

MANAGERIAL ECONOMICS

THEORY,
APPLICATION,
AND CASES

8TH | edition |

Allen

Weigelt

Doherty

Mansfield

MANAGERIAL ECONOMICS

EIGHTH EDITION

MANAGERIAL ECONOMICS

THEORY, APPLICATIONS, AND CASES

EIGHTH EDITION

W. Bruce Allen
The Wharton School
University of Pennsylvania

Neil A. Doherty
The Wharton School
University of Pennsylvania

Keith Weigelt
The Wharton School
University of Pennsylvania

Edwin Mansfield
late of University of Pennsylvania



W. W. NORTON & COMPANY
NEW YORK • LONDON

W. W. Norton & Company has been independent since its founding in 1923, when William Warder Norton and Mary D. Herter Norton first published lectures delivered at the People's Institute, the adult education division of New York City's Cooper Union. The firm soon expanded its program beyond the Institute, publishing books by celebrated academics from America and abroad. By midcentury, the two major pillars of Norton's publishing program—trade books and college texts—were firmly established. In the 1950s, the Norton family transferred control of the company to its employees, and today—with a staff of four hundred and a comparable number of trade, college, and professional titles published each year—W. W. Norton & Company stands as the largest and oldest publishing house owned wholly by its employees.

Copyright © 2013, 2009, 2005, 2002, 1999, 1996, 1993, 1990 by W. W. Norton & Company, Inc.

All rights reserved

Printed in the United States of America

Editor: Jack Repcheck

Editorial assistant: Hannah Bachman

Project editor: Kate Feighery

Senior production manager, College: Benjamin Reynolds

Art direction: Rubina Yeh

Media editor: Cassie del Pilar

Associate media editor: Nicole Sawa

Assistant media editor: Carson Russell

Marketing manager, Economics: John Kresse

Composition: Jouve International—Brattleboro, VT

Manufacturing: Transcontinental

Library of Congress Cataloging-in-Publication Data

Managerial economics: theory, applications, and cases/ W. Bruce Allen . . . [et al.]. – 8th ed.
p. cm.

Includes bibliographical references and index.

ISBN 978-0-393-91277-7 (hardcover)

1. Managerial economics. I. Allen, W. Bruce.

HD30.22.M354 2013

338.5024'658—dc23

2012018763

W. W. Norton & Company, Inc., 500 Fifth Avenue, New York, N.Y. 10110

W. W. Norton & Company, Ltd., Castle House, 75/76 Wells Street, London W1T 3QT

wwnorton.com

1 2 3 4 5 6 7 8 9 0

To Edwin Mansfield, a pioneer of managerial economics

BRIEF CONTENTS

PREFACE xix

PART 1: THE NEED FOR A GUIDE 1

CHAPTER 1 INTRODUCTION 2

PART 2: THE NATURE OF MARKETS 27

CHAPTER 2 DEMAND THEORY 28

CHAPTER 3 CONSUMER BEHAVIOR AND RATIONAL CHOICE 65

CHAPTER 4 ESTIMATING DEMAND FUNCTIONS 97

PART 3: PRODUCTION AND COST 135

CHAPTER 5 PRODUCTION THEORY 136

CHAPTER 6 THE ANALYSIS OF COSTS 172

PART 4: MARKET STRUCTURE AND SIMPLE PRICING STRATEGIES 223

CHAPTER 7 PERFECT COMPETITION 224

CHAPTER 8 MONOPOLY AND MONOPOLISTIC COMPETITION 257

PART 5: SOPHISTICATED MARKET PRICING 303

CHAPTER 9 MANAGERIAL USE OF PRICE DISCRIMINATION 304

CHAPTER 10 BUNDLING AND INTRAFIRM PRICING 357

PART 6: THE STRATEGIC WORLD OF MANAGERS 409

CHAPTER 11 OLIGOPOLY 410

CHAPTER 12 GAME THEORY 460

CHAPTER 13 AUCTIONS 501

PART 7: RISK, UNCERTAINTY, AND INCENTIVES 531

CHAPTER 14 RISK ANALYSIS 532

CHAPTER 15 PRINCIPAL-AGENT ISSUES AND MANAGERIAL
COMPENSATION 567

CHAPTER 16 ADVERSE SELECTION 619

**PART 8: GOVERNMENT ACTIONS AND
MANAGERIAL BEHAVIOR**

655

CHAPTER 17 GOVERNMENT AND BUSINESS 656

CHAPTER 18 OPTIMIZATION TECHNIQUES 728

**APPENDIX A TECHNOLOGICAL CHANGE AND
INDUSTRIAL INNOVATION 763**

**APPENDIX B BUSINESS AND ECONOMIC
FORECASTING 791**

APPENDIX C DISCOUNTING AND PRESENT VALUES 821

**APPENDIX D ANSWERS TO SELECT END-OF-
CHAPTER PROBLEMS 829**

APPENDIX E TABLES 861

INDEX 877

CONTENTS

PREFACE xix

PART 1: THE NEED FOR A GUIDE

1

CHAPTER 1 INTRODUCTION 2

- 3 The Theory of the Firm
- 5 What is Profit?
- 6 Reasons for the Existence of Profit
- 6 Managerial Interests and the Principal–Agent Problem
- 7 Demand and Supply: A First Look
- 11 The Demand Side of a Market
- 13 The Supply Side of a Market
- 14 Equilibrium Price
- 15 Actual Price
- 17 What If the Demand Curve Shifts?
- 18 What If the Supply Curve Shifts?
- 20 Summary
- 21 Problems
- 24 Excel Exercise: Demand, Supply, and Market Equilibrium

PART 2: THE NATURE OF MARKETS

27

CHAPTER 2 DEMAND THEORY 28

- 29 The Market Demand Curve
- 33 Industry and Firm Demand Functions
- 36 The Own-Price Elasticity of Demand
- 39 Point and Arc Elasticities
- 40 Using the Demand Function to Calculate the Price Elasticity of Demand
- 41 The Effect of Price Elasticity on the Firm's Revenue
- 43 Funding Public Transit
- 44 Determinants of the Own-Price Elasticity of Demand
- 45 The Strategic Use of the Price Elasticity of Demand

CONTENTS

48	Total Revenue, Marginal Revenue, and Price Elasticity
53	The Income Elasticity of Demand
55	Cross-Price Elasticities of Demand
57	The Advertising Elasticity of Demand
59	The Constant-Elasticity and Unitary Elastic Demand Function
60	Summary
61	Problems

CHAPTER 3 CONSUMER BEHAVIOR AND RATIONAL CHOICE 65

66	Indifference Curves
68	The Marginal Rate of Substitution
69	The Concept of Utility
70	The Budget Line
73	The Equilibrium Market Bundle
76	Maximizing Utility: A Closer Look
76	Corner Solutions
79	How Managers Can Strategically Influence Consumer Choices
84	Deriving the Individual Demand Curve
88	Deriving the Market Demand Curve
90	Consumer Surplus
92	Summary
93	Problems

CHAPTER 4 ESTIMATING DEMAND FUNCTIONS 97

97	The Identification Problem
100	Consumer Interviews
101	Market Experiments
101	Regression Analysis
103	Simple Regression Model
105	Sample Regression Line
106	Method of Least Squares
109	Coefficient of Determination
109	Multiple Regression
113	Software Packages and Computer Printouts
114	Interpreting the Output of Statistical Software
118	Multicollinearity
119	Serial Correlation
120	Further Analysis of the Residuals

123	Summary
125	Problems
131	Appendix: The Coefficient of Determination and the Concept of Explained Variation

PART 3: PRODUCTION AND COST

135

CHAPTER 5 PRODUCTION THEORY 136

137	The Production Function with One Variable Input
143	The Law of Diminishing Marginal Returns
143	The Production Function with Two Variable Inputs
146	Isoquants
149	The Marginal Rate of Technical Substitution
152	The Optimal Combination of Inputs
156	Corner Solutions
156	Returns to Scale
160	The Output Elasticity
161	Estimations of Production Functions
165	Summary
165	Problems
168	Appendix: Lagrangian Multipliers and Optimal Input Combinations

CHAPTER 6 THE ANALYSIS OF COSTS 172

173	Opportunity Costs
174	Short-Run Cost Functions
177	Average and Marginal Costs
184	Long-Run Cost Functions
189	Managerial Use of Scale Economies
191	Managerial Use of Scope Economies
194	Transactions Costs Can Take Many Forms
198	Network Economies
201	Managerial Use of Break-Even Analysis
203	Profit Contribution Analysis
204	Summary
205	Problems
209	Excel Exercise: Production and Cost
211	Appendix A: Break-Even Analysis and Operating Leverage
214	Appendix B: Measurement of Short-Run Cost Functions: The Choice of a Mathematical Form

**PART 4: MARKET STRUCTURE AND SIMPLE
PRICING STRATEGIES**

223

CHAPTER 7 PERFECT COMPETITION 224

- 225 Market Structure
- 227 Market Price in Perfect Competition
- 229 Shifts in Supply and Demand Curves
- 229 The Output Decision of a Perfectly Competitive Firm
- 234 Setting the Marginal Cost Equal to the Price
- 240 Another Way of Viewing the Price Equals Marginal Cost
Profit-Maximizing Rule
- 242 Producer Surplus in the Short Run
- 245 Long-Run Equilibrium of the Firm
- 247 The Long-Run Adjustment Process: A Constant-Cost Industry
- 249 The Long-Run Adjustment Process: An Increasing-Cost Industry
- 251 How a Perfectly Competitive Economy Allocates Resources
- 252 Summary
- 252 Problems
- 254 Excel Exercise: Perfect Competition

CHAPTER 8 MONOPOLY AND MONOPOLISTIC COMPETITION 257

- 259 Pricing and Output Decisions in Monopoly
- 267 Cost-Plus Pricing
- 269 Cost-Plus Pricing at Therma-Stent
- 269 Cost-Plus Pricing at Internet Companies and
Government-Regulated Industries
- 271 Can Cost-Plus Pricing Maximize Profit?
- 273 The Multiple-Product Firm: Demand Interrelationships
- 274 Pricing of Joint Products: Fixed Proportions
- 277 Output of Joint Products: Variable Proportions
- 281 Monopsony
- 282 Monopolistic Competition
- 285 Advertising Expenditures: A Simple Rule
- 287 Using Graphs to Help Determine Advertising Expenditure
- 288 Advertising, Price Elasticity, and Brand Equity: Evidence on
Managerial Behavior
- 289 Summary
- 290 Problems
- 294 Excel Exercise: Simple Monopoly
- 298 Appendix: Allocation of Output among Plants

PART 5: SOPHISTICATED MARKET PRICING**303****CHAPTER 9 MANAGERIAL USE OF PRICE DISCRIMINATION 304**

- 305 Motivation for Price Discrimination
- 306 Price Discrimination
- 318 Using Coupons and Rebates for Price Discrimination
- 319 Peak Load Pricing
- 324 Two-Part Tariffs
- 339 Summary
- 340 Problems
- 344 Excel Exercise: Perfect Price Discrimination
- 346 Excel Exercise: Third-Degree Price Discrimination
- 351 Appendix: Two-Part Tariff with Intersecting Demands

CHAPTER 10 BUNDLING AND INTRAFIRM PRICING 357

- 358 The Mechanics of Bundling
- 377 When to Unbundle
- 381 Bundling as a Preemptive Entry Strategy
- 386 Tying at IBM, Xerox, and Microsoft
- 388 Transfer Pricing
- 392 Transfer Pricing: A Perfectly Competitive Market for the Upstream Product
- 395 The Global Use of Transfer Pricing
- 398 Summary
- 399 Problems
- 403 Excel Exercise: Transfer Pricing

PART 6: THE STRATEGIC WORLD OF MANAGERS**409****CHAPTER 11 OLIGOPOLY 410**

- 411 Cooperative Behavior
- 413 The Breakdown of Collusive Agreements
- 414 Price Leadership
- 417 Possible Behavior in Markets with Few Rivals
- 435 Duopolists and Price Competition with Differentiated Products
- 439 The Sticky Pricing of Managers
- 440 Summary
- 441 Problems
- 447 Excel Exercise: Dominant Firm Price Leader

- 449 Excel Exercise: Cournot
- 457 Excel Exercise: Stackelberg

CHAPTER 12 GAME THEORY 460

- 460 Making Strategy and Game Theory
- 461 Strategy Basics
- 463 Visual Representation
- 467 Solution Concepts
- 467 Equilibria
- 468 Dominant Strategies
- 472 The Nash Equilibrium
- 479 Strategic Foresight: The Use of Backward Induction
- 483 Repeated Games
- 485 Incomplete Information Games
- 487 Reputation Building
- 489 Coordination Games
- 493 Strictly Competitive Games
- 494 Summary
- 495 Problems
- 498 Excel Exercise: Game Theory

CHAPTER 13 AUCTIONS 501

- 502 A Short History of Auctions
- 504 Types of Auction Mechanisms
- 505 Auction Mechanism and Revenue Generation
- 507 Bidding Strategies
- 510 Strategies for Sellers
- 515 Value of Information
- 518 Risk Aversion
- 519 Number of Bidders
- 520 Winner's Curse
- 522 Concerns in Auction Design
- 523 Summary
- 524 Problems
- 526 Excel Exercise: Auctions

PART 7: RISK, UNCERTAINTY, AND INCENTIVES

531

CHAPTER 14 RISK ANALYSIS 532

- 533 Risk and Probability
- 536 Probability Distributions and Expected Values

537	Comparisons of Expected Profit
538	Road Map to Decision
540	The Expected Value of Perfect Information
542	Measuring Attitudes toward Risk: The Utility Approach
548	Attitudes toward Risk: Three Types
551	The Standard Deviation and Coefficient of Variation: Measures of Risk
553	Adjusting the Valuation Model for Risk
554	Certainty Equivalence and the Market for Insurance
557	Summary
558	Problems
562	Excel Exercise: Expected Utility

CHAPTER 15 PRINCIPAL–AGENT ISSUES AND MANAGERIAL COMPENSATION 567

567	Principal–Agent Issues
569	The Diverging Paths of Owners and Managers
570	The Principal–Agent Situation
572	The Effect of Risk, Information, and Compensation on Principal–Agent Issues
581	Resolving the Incentive Conflict When Output Is Risky and Effort Is Not Observable
589	Some Refinements to Managerial Compensation
594	Principal–Agent Issues in Other Contexts
607	Product Liability and the Safety of Consumer Goods
612	Summary
613	Problems
616	Excel Exercise: Moral Hazard

CHAPTER 16 ADVERSE SELECTION 619

620	The Market for “Lemons”
622	Adverse Selection in Automobile Insurance
627	The Market for Annuities
634	Resolving Adverse Selection through Self-Selection
639	Using Education as a Signal: Adverse Selection in the Job Market
640	Using Warranties as Signals: Adverse Selection in the Product Market
648	Summary
648	Problems
650	Excel Exercise: Adverse Selection

PART 8: GOVERNMENT ACTIONS AND MANAGERIAL BEHAVIOR 655

CHAPTER 17 GOVERNMENT AND BUSINESS 656

- 658 Competition versus Monopoly
- 660 Regulation of Monopoly
- 664 The One Star Gas Company: A Pseudo-Case Study
- 666 Effects of Regulation on Efficiency
- 667 The Concentration of Economic Power
- 674 The Sherman Act
- 676 The Clayton Act, the Robinson-Patman Act, and
the Federal Trade Commission Act
- 677 Interpretation of the Antitrust Laws
- 680 The Patent System
- 682 Trade and Trade Policy
- 698 Government Price Ceilings and Price Floors
- 702 The Welfare Impacts of Taxes
- 703 Regulation of Environmental Pollution
- 714 Public Goods
- 720 Summary
- 721 Problems
- 725 Excel Exercise: Externalities

CHAPTER 18 OPTIMIZATION TECHNIQUES 728

- 728 Functional Relationships
- 729 Marginal Analysis
- 731 Relationships among Total, Marginal, and Average Values
- 734 The Concept of a Derivative
- 738 How to Find a Derivative
- 744 Using Derivatives to Solve Maximization and
Minimization Problems
- 748 Marginal Cost Equals Marginal Revenue and the Calculus
of Optimization
- 750 Partial Differentiation and the Maximization of
Multivariable Functions
- 752 Constrained Optimization
- 754 Lagrangian Multipliers
- 756 Comparing Incremental Costs with Incremental Revenues
- 758 Summary
- 758 Problems

APPENDIX A TECHNOLOGICAL CHANGE AND INDUSTRIAL INNOVATION 763

763	Technological Change
765	Labor Productivity
766	Total Factor Productivity
767	Using Total Factor Productivity to Track Factory Performance
769	Research and Development: A Learning Process
770	Parallel Development Efforts
771	What Makes for Success?
772	Project Selection
774	Innovation
774	Time-Cost Trade-Offs
777	The Learning Curve
777	Applications of the Learning Curve
779	Henry Ford's Model T and Douglas Aircraft's DC-9
780	Diffusion Models
782	Forecasting the Rate of Diffusion of Numerically Controlled Machine Tools
785	Summary
786	Problems

APPENDIX B BUSINESS AND ECONOMIC FORECASTING 791

792	Survey Techniques
793	Taking Apart a Time Series
795	How to Estimate a Linear Trend
797	How to Estimate a Nonlinear Trend
798	Seasonal Variation
799	Calculation of Seasonal Variation
802	Cyclical Variation
804	Elementary Forecasting Techniques
806	How Leading Indicators are Used
807	How Econometric Models are Used
809	The Purvere Corporation: A Numerical Example
810	"Study Your Residuals"
812	Summary
813	Problems
817	Appendix: Exponential Smoothing and Forecasting

CONTENTS

APPENDIX C DISCOUNTING AND PRESENT VALUES 821

- 823 Present Value of a Series of Payments
- 824 The Use of Periods Other Than a Year
- 826 Determining the Internal Rate of Return

APPENDIX D ANSWERS TO SELECT END-OF-CHAPTER PROBLEMS 829

APPENDIX E TABLES 861

INDEX 877

PREFACE

Where does the time go? Four years ago we wrote this preface in a world turned upside down. Akin to natural disasters, deep global recessions can have devastating effects. The U.S. economy is finally on the brink of relative recovery, though housing prices are still scrapping bottom and unemployment remains stubbornly high.

The recession reinforced many lessons: that unbridled business greed can cause misery for billions; that politicians, though agents of the people, are equally selfish when they behave as if political ideology is paramount to the collective good; and, finally, that the human spirit is difficult to extinguish even when overwhelmed by financial ruin or the army of a tyrant.

The recession gave us the opportunity to test the relevancy of our book under extreme business conditions. The fact that we did not change the principles discussed in this edition shows our satisfaction with the decision framework. In fact, the recession actually made us popular party guests given our knowledge of moral hazard, one of the key topics presented in the text. Prior to the recession any party discussion of moral hazard would have been met with yawns. After the financial crash, guests would hang on our every word regarding it.

We have structured the text on the principle that managers need to use economics, the language of business, to make decisions. Students need to see past the equations and graphs, and understand how to follow certain principles. For example, managers need to always use marginal analysis and backwardly induct.

The formal analysis of economic models adds value to managerial discussions. Without it, perceived knowledge is mere speculation. When developing incentive schemes managers prefer those with objective metrics. We believe the objectivity of mathematics brings focus to business decisions. Its mastery is essential for good managerial decision making.

We recognize that managers are inherently rational; however they occasionally make mistakes in judgment. Understanding this book will help managers avoid common mistakes. Managers make decisions in a complex world, so carefully thinking through alternatives is important. This is especially true since business decisions can involve potential losses of billions of dollars.

THE EIGHTH EDITION OF MANAGERIAL ECONOMICS

Because the managerial world is one of change, we focused on revising the text to make it more relevant to our current world. In so doing, we have listened to the suggestions of readers. They compelled us to include discussions of network externalities and transaction costs in this edition. We also returned chapters on estimating demand functions and forecasting based on the comments of readers. Finally, we updated many of our cases to illustrate the relevancy of our book to the changed business environment. Many of our new cases are not U.S. centric. Business is truly global and so are the principles managers should follow. Many of our new cases are based on emerging markets. We believe these markets will exhibit the fastest growth in the near future. Managers need to understand them well. Below are brief summaries of new material relative to the seventh edition.

Chapter 1: Introduction

We updated the efforts of Bono to get others to donate money to the RED campaign: as of the beginning of 2011, over 55 billionaires have pledged to contribute at least 50% of their wealth to charitable causes.

We updated the progress of Major League Baseball Teams to adjust their ticketing plans to market conditions. Gratifying to us, some teams are following policies that we suggested in our last edition.

In two new Strategy Session boxes, we show how supply and demand determine market prices across numerous markets. We show the invisible hand in markets for pineapples, oil, beef, baseball, and stocks. In the stock market, we show how shifts in demand and supply impact the effectiveness of hedging strategies. The range of examples illustrates the pervasive power of using demand and supply shifts to explain market behavior.

Chapter 2: Demand Theory

We updated the Strategy Session box on how managers use dynamic pricing models to increase margins through the setting of better prices. This practice continues to grow as predicted in past editions. We also updated the business fortunes of Amtrak managers, especially in their Northeast Corridor business. We identify primary causes for shifts in demand.

In a new Strategy Session box, we investigate the demand for beer, a subject most college students can relate to. We find the biggest consumers of beer, young males, are fickle yet rational lovers. When their incomes drop enough, even they must reject their love of beer.

Another new Strategy Session box looks at the suitability of using direct marketing to reach women in developing economies. Women in these markets appear to value the personal relationship they establish with the sales representative.

We also report on the actions of pharmaceutical managers in changing their pricing structure for a specialty drug. When threatened with an entry of an alternative drug, the managers actually increased their prices. We show they did so because of inelastic demand.

Chapter 5: Production Theory

To show how costs affect the mix of capital and labor, we note how artificial intelligence software is replacing labor in the legal market. The software is so effective that one lawyer with the software does the work of one hundred lawyers without it.

We look at the increase in CIOs (Chief Innovation Officers). CIOs help firms adapt to new technologies, like touch-screen workstations. These technologies generally replace labor, which results in either a smaller workforce or a shifting of responsibilities.

We also report on the use of smartphones as mobile banking begins to transform the cash society. All major players in the mobile eco-system—telecom operators, banks, search engines, and hardware manufacturers—are racing to gain control of the market.

Baseball teams are a type of business organization. We find empirical support that teams do better in the playoffs when they possess what are called “power pitchers.” The tighter schedule of playoff games relative to games in the regular season means that this type of pitcher is more effective.

Chapter 6: The Analysis of Costs

Managers often generate cost efficiencies through economies of scope. We discuss how managers generate these cost savings, and also discuss the difficulties that they encounter.

Chapter 7: Perfect Competition

We showcase many markets where competition tends to drive the price down to marginal cost. Markets that we look at include power transmission, iron ore, mining, and cement.

Chapter 8: Monopoly and Monopolistic Competition

Who better to decide the relevancy of our book than Warren Buffett? We use his actions to illustrate the effectiveness of our suggested strategies throughout the book. In this Strategy Session box we discuss why Mr. Buffett thinks that pricing power is the most important criteria to examine when evaluating a business.

By creating the market for tablets, Apple managers have done something very few managers accomplish. We show how these managers use their pricing power when selling the tablets.

There is some evidence that monopolists can charge a price that seems excessively high. We discuss one such situation of a railroad in South Africa where managers increased their freight rates over 25% in one year.

Chapter 9: Managerial Use of Price Discrimination

We add another example of how managers use price discrimination to increase profits. We tell the story of how several major universities use this strategy to increase their profits from sporting events.

Chapter 10: Bundling and Intrafirm Pricing

One common complaint of students is the high price of textbooks at their campus book store. We discuss how managers tried to increase the price by bundling the text books with study guides. This forced students to buy both. In response, several state legislatures impose laws requiring book stores to change their policies.

Managers of Netflix used a very successful pricing strategy that allowed customers to either rent DVDs of movies or directly stream them to their computers. In the quest for higher profits they decided to change the strategy. We discuss why millions of users deserted them because of the change.

We update a past discussion on the pricing of printers and their ink cartridges. We also discuss a recent strategy by one company to change their pricing strategy in the market.

We show the many ways that managers use transfer pricing to achieve their goals. The focus is on the global aspects of the strategy. We also report on a 2010 study by Ernst and Young.

Chapter 11: Oligopoly

We report on one way Warren Buffett earns his billions—by investing in a duopoly. Yes, they do actually exist, and Buffett is smart enough to realize their profit potential. The duopoly in the civil jet industry of Airbus and Boeing is under attack by new entrants. Many of these entrants reside in emerging markets, like Brazil and China.

Our final new addition to the chapter shows how many markets across the globe resemble oligopolies. These range from audit firms in Europe to rating agencies in the United States. Even in emerging markets one can find oligopolies, such as the telephone industry in the Philippines.

Chapter 13: Auctions

As we predicted in the last edition, the use of auctions in the business world continues to increase. We show several examples of its increased use in sports and the pricing of ships using the Panama Canal.

Coke and Pepsi are iconic American brands. They have been battling for market share against one another for decades. We discuss how this focus on market share has resulted in lower profitability for both firms.

Chapter 15: Principal–Agent Issues and Managerial Compensation

Private equity firms often purchase firms and hold them for short periods of time. In that period, they re-structure assets and then sell them, often earning significant profits. Frequently, much of this profit is earned by selling off corporate assets like corporate jets and city apartments that prior managers purchased with shareholder money. We report on several examples.

It is said that European managers earn lower compensation relative to their U.S. counterparts. Using empirical data we show that this is true, though the gap is getting smaller.

We report on a smartphone app that retailers can use to track the location of customers. Through the app they can offer discounts to lure shoppers as they pass a local outlet of the retailer.

Safaricom is the largest telecom operator in Kenya. We describe the actions of their managers as they cope with new entrants.

All emerging markets must follow an evolution in their antitrust laws. We report on the recent efforts of the Chinese government to mitigate the power of dominant market players.

The Europeans actually take a harder look at anti-competitive behavior than their counterparts in the United States. We report on actions taken against a pharmaceutical company that exercised too much market power.

We discuss an interesting case where a food product that could help billions of the world's poor is not used as extensively as it could because of patents on its underlying intellectual property. Some question whether patents should hamper such a clear social good.

Finally, we report how managers in pharmaceutical companies try to protect their patents by slightly changing the chemical formula of a drug. Such actions may help them protect their high price (and market share) for longer periods of time.

ANCILLARIES FOR STUDENTS AND INSTRUCTORS

For Instructors

Instructor Site (wwnorton.com/instructors): the instructor only, password-protected site features instructional content for use in lecture and distance education, including coursepacks, test-item files, PowerPoint lecture slides, figures, and more.

Instructor's Manual: paperback and downloadable as PDF; contains all of the solutions to the end-of-chapter problems, answers to suggested discussion questions for the Strategy Sessions/Case Studies, plus detailed outlines.

Test Bank: paperback and downloadable.

PowerPoints: for all of the figures and tables from the text.

For Students

The student media package for *Managerial Economics*, Eighth Edition, provides students with the tools to master the key concepts and skills covered in the text, helps students gain familiarity with the essential tools used in the business world, and gives them opportunities to understand how the economic lessons from the text translate into real business-world situations.

StudySpace: Your Place for a Better Grade (www.norton.com/studyspace)

StudySpace tells students what they know, shows them what they still need to review, and then gives them an organized study plan to master the material. This free and easy-to-navigate website offers students an impressive range of exercises, interactive learning tools, assessment, and review materials, including:

- Review quizzes
- Chapter summaries
- Chapter outlines
- Flashcards
- Updated Excel exercises
- Audio Podcasts from MarketPlace Business Reports

A Note of Thanks

Finally, we would like to thank the hundreds of instructors who have used our text over the years, and the many who have provided valuable feedback for improving it. For help in preparing for this Eighth Edition, we want to specifically acknowledge James Ahiakpor (Cal State University, East Bay), Hamid Bastin (Shippensburg University), Bichaka Fayissa (Middle Tennessee State University), Sherman Folland (Oakland University), Chulho Jung (Ohio University), Barry Keating (University of Notre Dame), Mark Moore (Simon Fraser University), Kamal Salmasi (McGill University), Manuel Santos (University of Miami), and Ryan B. Williams (Texas Tech University). We would also like to thank David Muir and Vivek Shah, two outstanding Ph.D. students here at The Wharton School, who helped us to make sure that all of the math in the text is as clear as possible. Two outstanding instructors have assisted us on the ancillaries and a simple thank you is not nearly enough—Nitin Paranjpe (Oakland University and Wayne State University) and Jean Cupidon (Texas Tech University). Last and certainly not least, we have been fortunate to have a real star at our publisher, W. W. Norton, who transformed our manuscript into a cohesive text—Kate Feighery.

PART 1

THE NEED FOR A GUIDE

CHAPTER 1

LEARNING OBJECTIVES

The Theory of the Firm
What Is Profit?
Reasons for the Existence
of Profit
Managerial Interests and the
Principal-Agent Problem
Demand and Supply:
A First Look
The Demand Side of a Market
The Supply Side of a Market
Equilibrium Price
Actual Price
What If the Demand
Curve Shifts?
What If the Supply Curve Shifts?
Summary
Problems
Excel Exercise: Demand, Supply,
and Market Equilibrium

INTRODUCTION

The main task of managers is to make good decisions. For better or worse, managers face a complex world, and they need a guide to help them choose well. This is that guide. Those who gain its understanding will increase the value of their decisions at personal and organizational levels.

This guide provides knowledge in the following sense. The ancient Chinese discuss knowledge as a temporal flow. Knowledge is not storage of memorized facts but an ability to understand the actions of others. With this knowledge, you better anticipate their behavior. Our guide will help you navigate through the managerial world of behavior.

We construct our guide within the framework of managerial economics. Managerial economics uses formal models to analyze managerial actions and their effect on firm performance. We use these models to shed light on business concepts such as cost, demand, profit, competition, pricing, compensation, market entry strategy, and auction strategy. All these concepts are under the control of managers, and they determine firm performance.

Contrary to the beliefs of many, managerial economics differs significantly from microeconomics: The focus of analysis is different. At best, the focus in microeconomics is at the firm level; many times the analysis is at the market level. In managerial economics, the focus is on managerial behavior. Managerial

economics prescribes behavior, whereas the micro world describes the environment. This focus on managerial behavior provides powerful tools and frameworks to guide managers to better decisions. These tools allow managers to better identify the consequences of alternative courses of action.

Managerial economics plays two important roles in preparing students for managerial life. Concepts we will discuss in subsequent chapters are found in other functional business courses like accounting, finance, strategy, operations, and marketing. Our guide is what the great strategist Sun Tzu called the “general’s seat,” and it is characterized by what are known as *economies of scope*. That is, the better you understand the concepts we discuss, the easier will be your understanding of them when they arise in other business classes. And because managerial economics recognizes the complexity of the managerial world, it is arguably the most integrative of the functional business classes. This helps students learn the integrative mind-set that is essential for good management, and it also gets them to think past the short-term mentality and consider the long-term consequences.

THE THEORY OF THE FIRM

Managers work within a larger organization and ultimately determine its performance. To understand the behavioral world of managers, we must account for the behavior of firms. Of course, firms really don’t behave on their own; you might think of them as marionettes with managers controlling the strings. Some management teams are good at pulling these strings, while others can’t seem to get it right. But although management styles differ greatly in the millions of firms across the globe, there is surprisingly little variance in the goals of managers. Overwhelmingly, managers choose actions they believe will increase the value of their organization. So in our theory of the firm, the goals of managers focus on increasing this value. We understand there are many ways to create value in an organization; for example, to a microcredit organization with a double bottom line, value from its lending practices might consist of a profit measure and the gains to a local community’s economy. But our models must account for behavior across a great number of firms, so we take the view that managers in profit-oriented organizations try to increase the net present value of expected future cash flows. We can formally present this managerial effort in the following

$$\begin{aligned} \text{Present value of expected future profits} &= \frac{\pi_1}{1+i} + \frac{\pi_2}{(1+i)^2} + \cdots + \frac{\pi_n}{(1+i)^n} \\ \text{Present value of expected future profits} &= \sum_{t=1}^n \frac{\pi_t}{(1+i)^t} \end{aligned} \quad (1.1)$$

STRATEGY SESSION: Bono Sees Red, and Corporate Profits See Black

In 2007, rock star Bono started Red, a campaign that combines consumerism with altruism. When a consumer buys a Dell Red computer, a Motorola Red Motorazr, or Red items from the GAP, or uses an American Express Red card, a contribution is made into the global fund. Companies pay a licensing fee to label their products “Red” and then pay a portion of the sales from those products into the fund. According to the Red website (www.joinred.com/red/), \$170 million has been generated by the Red campaign and been put to its intended use fighting AIDS, malaria, and tuberculosis in the African countries of Ghana, Rwanda, and Swaziland.

This concept has been branded *cause marketing* and was around long before Red—but without the publicity generated by a personality like Bono. For years McDonald’s has sponsored Ronald McDonald Houses, where parents of sick children can find respite (and support groups) near the hospitals where their children are patients. Why do firms participate in cause marketing? Clearly, the companies can market themselves as being socially conscious. But equally clear is the cost of paying for the Red license and operating Ronald McDonald Houses. Where’s the tangible benefit? According to a 2006 poll by Cone Inc. (a Bos-

ton marketing agency), 89% of Americans aged 13 to 25 (a large consumer group and one swayed by Bono) would switch to a brand associated with a “good cause” if the products and prices were comparable. There’s the tangible benefit: Cause marketing also leads to a revenue stream (and presumably a profit stream). So incurring Red leads to seeing black as profits increase because of such actions. Cause marketing is in harmony with profit maximization.

Another phenomenon we are witnessing is the rise of “philantreneurs” such as Bill Gates, Warren Buffett, Richard Branson, and Ted Turner—individuals whose businesses make a lot of profit and wealth for themselves and then give a lot of that money to good causes. As of December 2010, in the United States, 57 billionaires have pledged to give a minimum of 50% of their wealth to good causes under something known as the “Giving Pledge.” In addition to the billionaires named above, Paul G. Allen, Michael R. Bloomberg, Larry Ellison, Jon Huntsman, Carl Icahn, George Lucas, and Mark Zuckerberg are among those taking the pledge.

Source: “Bottom Line for [Red],” *The New York Times*, February 6, 2008; The Giving Pledge, www.givingpledge.org.

where π_t is the expected profit in year t , i is the interest rate, and t goes from 1 (next year) to n (the last year in the planning horizon). Because profit equals total revenue (TR) minus total cost (TC), this equation is also expressed as

$$\text{Present value of expected future profits} = \sum_{t=1}^n \frac{TR_t - TC_t}{(1 + i)^t} \quad (1.2)$$

where TR_t is the firm’s total revenue in year t , and TC_t is its total cost in year t .

Equation (1.2) shows why managers influence firm performance. Managerial decisions clearly determine both the revenues and costs for an organization. Consider, for example, the Toyota Motor Company. Its marketing managers and

sales representatives work hard to increase its total revenues, while its production managers and manufacturing engineers strive to reduce its total costs. At the same time, its financial managers play a major role in obtaining capital and hence influence equation (1.2); its research and development personnel invent new products and processes to increase total revenues and reduce total costs. Managers of all these diverse groups make decisions to affect Toyota Motor's value, defined here as the present value of expected future profits.

Although managers want to increase their firm's value, they do not have total control over the level of value. If managerial life were that simple, you would not have to go to school to learn business techniques. What complicates managerial life are the operating constraints managers face. One constraint is that most resources are scarce. Within the firm, managerial decision making often involves allocating scarce inputs to support the production, distribution, and sales of goods and services that are sold at a price that exceeds their costs.

Other constraints that limit managerial actions are legal or contractual. For example, managers may be bound to pay wages exceeding a certain level because minimum wage laws stipulate that they must do so. Also, they must pay taxes in accord with federal, state, and local laws. Further, managers must comply with contracts with customers and suppliers—or take the legal consequences. A wide variety of laws (ranging from environmental laws to antitrust laws to tax laws) limit what managers can do, and contracts and other legal agreements further constrain their actions.

WHAT IS PROFIT?

As we have seen, firm value is largely a function of **profit**. Unlike in accounting, in managerial economics we measure profit after taking account of the capital and labor provided by the owners. For example, suppose a manager quits her position at a large firm to create a small start-up business. She receives no salary even though she puts in long hours trying to establish her business. If she worked these hours for her previous firm, she would have earned \$65,000. And if she had invested the capital she used to begin her business in some alternative investment, she could have earned \$24,000. Let's say in 2012 her start-up firm earned an accounting profit of \$100,000. Her firm's profit in the managerial economics world is $\$100,000 - \$65,000 - \$24,000 = \$11,000$ rather than the \$100,000 shown in accounting statements.

The differences between the profit concepts used by the accountant and the economist reflect a difference in focus. The accountant is concerned with controlling the firm's day-to-day operations, detecting fraud or embezzlement, satisfying tax and other laws, and producing records for various interested groups. The economist is concerned with decision making and rational choice among

Profit When economists speak of profit, they mean profit over and above what the owner's labor and capital employed in the business could earn elsewhere.

strategies. Although most of a firm's financial statements conform to the accountant's and not the managerial economist's concept of profit, the latter is more relevant for managerial decisions. (And this, of course, is recognized by sophisticated accountants.) For example, suppose the woman is trying to decide whether to continue operating her business. If she is interested in making as much money as possible, she should calculate her firm's profit based on our economist model. If the firm's economic profit is greater than zero, she should continue to operate the firm; otherwise she should close it and pursue other opportunities.

REASONS FOR THE EXISTENCE OF PROFIT

A firm's economic profit is generated by the actions of managers. Profit is one indicator of their decision-making skills. Three fertile profit-generating areas used by managers are innovation, risk, and market power. As we write this chapter, people are waiting for the chance to buy the iPhone 4S, the new model of Apple's iconic smartphones. And airlines are committing billions of dollars for the opportunity to purchase the 787 Dreamliner from Boeing. In both these markets, products already exist; but consumers apparently are more interested in new products. Both the iPhone and the 787 are considered pioneering products. They push the frontier relative to existing products in terms of functionality, technology, and style. As we write today, these managerial efforts both generate high profit—reportedly up to 40%. Future value depends on how each managerial team executes its strategy.

A hallmark of managerial decision making is the need to make risky choices. For managers this risk takes many forms. They are asked to make decisions whose future outcomes are unknown (How successful will this product be in the market?), when they don't know the reactions of rivals (If I raise my price, will my rivals raise theirs?), and when they do not know the likelihood of a future event (How likely is it a Democrat is elected our next president?). Profit is the reward to those who bear risk well.

As we will see later, managers also earn profit by exploiting market inefficiencies. Good managers understand how to create these to give their firm a sustainable competitive advantage. Common tactics in this area include building market entry barriers, sophisticated pricing strategies, diversification efforts, and output decisions. Such tactics, if done well, can generate a long stream of profit.

MANAGERIAL INTERESTS AND THE PRINCIPAL-AGENT PROBLEM

Although managerial economists generally assume that managers want to maximize profit (and hence firm value, as defined in equation (1.1)), they recognize additional goals. Some goals may enhance the firm's long-term value, like building

market share or establishing a brand name. Other managerial goals have less to do with firm value and more to do with increasing managerial compensation.

As we will see, our model recognizes preferences of firm owners and managers sometimes diverge. And when managers make choices between maximizing a firm's value and increasing the payoffs to a single manager or management team, some choose the selfish path. This too is a trait of managerial behavior. The tendency to focus on self-interest is growing in importance because the separation between the ownership and management of firms is continuing to increase on a global scale. The owners of the firm—the stockholders—usually have little detailed knowledge of the firm's operations. Even a firm's board of directors has limited information relative to the management team. Managers are generally given a great deal of freedom as long as they perform adequately. Consequently, firm behavior is often driven by the interests of the nonowner management group. At the least, this behavior results in higher pay and more perquisites for managers; at worst, it creates an Enron spectacle.

Managerial economists call this the **principal-agent problem**. Managers are agents who work for the firm's owners, who are shareholders or principals. The principal-agent problem centers on whether managers may pursue their own objectives at a cost to the owners. We ask students in our class, "If we send you to Atlantic City with our money, would your behavior change?" Because the firm's owners find it difficult to adequately distinguish actions that maximize profit and those that do not, managers have incentives to enrich themselves.

To deal with this problem, owners often use contracts to converge their preferences and those of their agents. For example, owners may give managers a financial stake in future success. Many corporations use stock option plans, whereby managers can purchase shares of common stock at less than market price. These plans give managers incentives to increase firm profit and comply with owners' interest. There is some evidence these plans do change behavior. According to one study, if managers own between 5 and 20% of a firm, they are likely to perform better (that is, earn more profit) than if they own less than 5%. In some firms managers are forced to purchase stock, and boards of directors are compensated in stock. This and other moral hazard issues are discussed extensively in Chapter 15.

The principal-agent problem
When managers pursue their own objectives, even though this decreases the profit of the owners.

DEMAND AND SUPPLY: A FIRST LOOK

To understand behavior in any society, we must have a working knowledge of its institutions. The managerial world revolves around markets. Any manager, whether in Tokyo, New York, London, or Toronto, must understand basic market principles in order to anticipate behavior. A significant portion of this book is devoted to helping you understand the behavior of people in markets. We first give an overview of markets and then examine both the demand and supply sides in greater detail.

STRATEGY SESSION: Baseball Discovers the Law of Supply and Demand

It started several years ago with the Colorado Rockies looking for a way to obtain more revenue but at the same time not wishing to heap additional expense on their loyal season-ticket holders (who buy tickets for every game or an aggregation of games as a bundle). As of the 2003 season, the Rockies were joined by 11 other teams—slightly less than half of the 30 major league baseball teams. Doing what? Practicing what they call *variable pricing*. In lay terms, they are charging more for the exact same seat for games with desirable teams as opponents, such as teams that are traditional rivals, teams with superstars, and the like. And the trend continues. Almost every team now practices some sort of variable pricing.

Although this practice is nothing new for many goods and services (Miami Beach hotel rooms cost more in February than in July, ticket prices for Lady Gaga in the same seat in the same arena are more expensive than ticket prices for Carrie Underwood, etc.), it was new for baseball.

Historically, the price for seat X in the stadium was price Y for each of the team's 81 home games. Now seat X can be priced higher on opening day, on fireworks night, or when the New York Yankees or a traditional rival comes to town. The basic premise is the law of supply and demand. The number of seats in the ballpark remains fixed, but the attractiveness of the seat to a potential buyer is not constant.

Let's see what some of the teams are doing. Several teams, e.g., the Chicago Cubs, the New York Mets, the Tampa Bay Rays, and the Kansas City Royals, divide the games up into categories of perceived attractiveness. The Cubs have five categories (with the number of games in each category in parenthesis): Marquee (13), Platinum (13), Gold (31), Silver (13), and Bronze (11). The per-game price for the highest class seat (an Infield Club Box) for a Marquee game is \$112, while the same seat for a bronze game is \$58 (a 93% difference). The per-game price for the

lowest class seat (Reserved Outfield Upper Deck) is \$27 for a Marquee game and \$8 for a Bronze game (a 237.5% difference). The Mets have four categories: Marquee (four-against the Yankees), Premium (22), Classic (25), and Value (30). Their highest class seats are Delta Club Platinum and Gold, and their price remained the same (\$440 and \$325) through all four categories. But their third highest class seat (Delta Club Silver) was priced at \$280 for a Marquee game and \$160 for a Value game (a 75% difference). Cheaper ticket classes followed the same 75% difference that exists in the Marquee to Value games. The Rays also have four categories: Diamond, Platinum, Gold, and Silver. Their highest class seat (Avantair Home Plate Club) sells for \$300 for Diamond games and \$210 for Silver Games (a 43% difference). Their cheapest class (Upper Reserved Party Deck) sells for \$19 for Diamond games and \$9 for Silver Games (an 111% difference). The Royals have two categories: Premium and Regular. The highest class seat (BATS Crown Club) sells for \$250 for Premium games and \$240 for Regular games (a 4.2% difference) and their lowest class seat (Hy-Vee View) sells for \$15 for Premium games and \$10 for Regular Games (a 50% difference). The Royals also charge an additional \$1 or \$2 fee for tickets purchased on the day of the game.

The St. Louis Cardinals price are based on day of the week and the opponent faced. For instance, for July 2011 Monday, Tuesday, and Wednesday games against the Reds and the Astros, the highest class seats (Diamond Box) was \$110 for both, but Thursday games against the Diamondbacks and the Astros were \$101 and \$110 respectively. But Diamondback tickets for the Friday and Sunday games were \$120, and the Saturday game against the Diamondbacks was \$143. When the rival Cubs arrive two weeks later, the Friday and Saturday games are \$158 and the Sunday game \$130. Their cheaper seat classes follow a similar pattern.

The Atlanta Braves price based on the day of the week, regardless of the opposition faced. Monday through Thursday games are called Regular. Friday, Saturday, and July 4 are called Premium. Saturday is called Saturday and is their highest price category. Four of their 12 classes of seats have a fourth category (Value Days). Their highest seat class (Hank Aaron Seats) sells for \$90 on Saturdays but \$78 on Regular days (a 15.4% difference) and the lowest seat class (Upper Pavilion) sells for \$8 regardless of the day. When the Value tickets are available, there's a 35–50% difference between them and the Regular tickets.

So some teams differentiate prices based on the opposition, some based on the day of the week, and others based on both.

We've reported on Major League Baseball ticket prices in boxes like this in previous editions of this book and on how teams use *pre-season* information on perceived demand for games to determine prices. But this is subject to a lot of uncertainty. While a rival will still be a rival, a pre-season predicted strong draw may not pan out (because the team does not live up to expectations, because superstars get injured or have off seasons, the weather may turn bad, etc.). When teams post their prices, they have to live with them.

However, we also stated that "true variable pricing will see price changing very close to game time." That, of course, has always taken place in the illegal selling of tickets by sleazy characters hanging out in the vicinity of stadiums and with legitimate resellers of tickets like StubHub. But now the San Francisco Giants have taken the concept to a whole new level.

It started in 2010 when the Giants noticed a surge in ticket sales for the Memorial Day game with the Colorado Rockies. Normally, a Memorial Day game is a tough sell, as families opt for parades and picnics rather than baseball games. But that Memorial Day, tickets were selling like hotcakes. Why? Because the

Giants' pitching ace, Tim Lincecum, was facing the Rockies' pitching ace, Ubaldo Jimenez. So the Giants decided to raise ticket price for the game. For example, the Giants raised some ticket prices from \$17 to up to \$25. Even with such increases, the Giants sold 10,000 tickets on the weekend leading up to the Monday game, and the stadium was sold out for the game.

What the Giants noticed was that factors *during* the season, such as, weather, winning streaks, pitching matchups, and so on, that are only known as game time approaches can appreciably change the demand to see the game. Because the demand to see the game changes, the teams are now asking: Why shouldn't the game prices change? So the Giants are now changing their ticket prices *daily* depending on market conditions. The pricing is now almost truly dynamic (truly dynamic would be minute-to-minute changes as practiced by the scalpers).

The Giants didn't do this without experimenting. In 2009, they used the 2,000 least desirable seats in the stadium and raised or lowered prices based on their estimates of demand to see a game. They estimate that they sold 25,000 extra tickets yielding a gross revenue increase of \$500,000. In 2010, pricing all tickets dynamically has led to an estimated revenue increase of 6%.

"We debated the merits of dynamic pricing for years, but there were numerous hurdles," says Giants CIO Bill Schlough, such as protecting season-ticket holders and choosing the right variables for the pricing model. And here's where it's wonderful to be an economist. Schlough goes on to say "Dynamic pricing enables us to capture a larger share of consumer surplus . . ." (See the definition of *consumer surplus* on p. 90).

Won't some fans complain about higher ticket prices? Sure. When the Giants always charged X for seat Y, fans who were willing to pay X or more bought tickets, and fans willing to pay less than X did not.

(continued)

STRATEGY SESSION: Baseball Discovers the Law of Supply and Demand (continued)

Now, when the Giants charge $X + Z_1$ for the ticket, a number of fans still buy tickets (because they value seeing the game at $X + Z_1$ or greater). So some fans don't go to games anymore (those who value seeing the game at slightly less than $X + Z_1$ and X), and those that do attend pay more to the Giants. Neither of those types is likely to be overjoyed, but when the Giants charge $X - Z_2$ for a ticket, the people willing to pay X (or more) will purchase tickets, as will fans who have never purchased tickets before. This can lead to an expansion of the fan base. The Giants, being rational, wouldn't price this way unless it increased profits.

Like airline yield management models, sports teams beyond baseball, such as basketball and hockey, are using computers to analyze past ticket sales and data from the above mentioned legal secondary markets to try to price games based on what the markets will bear.

The models have many independent variables, and the game ticket price is the variable to be predicted. The models generate weights associated with each variable, such as day of the week, team perfor-

mance, pitching matchups, and so on, to yield a suggested ticket price for the seat for the game. Then a human being takes the recommended price and decides whether to adjust it and by how much.

The Giants make pricing decisions on future games *every morning*. So, as with airline fares, a price for a game two weeks from now could change daily up to game day. Fans seem to be accepting the process because they have become accustomed to it in the secondary market for tickets.

In the NBA (basketball), more than half the teams used a model to set season-ticket prices last year. A third of teams will use per-game pricing this year. In Cleveland (which had superstar LeBron James at the time), the Cavaliers priced 20,000 tickets on a daily basis and reported an average price increase of \$9.25 per ticket.

Source: Various teams' websites; "Star Pitchers in a Duel? Tickets Will Cost More," *The New York Times*, June 28, 2010, at www.nytimes.com/2010/06/28/technology/28tickets.html; "San Francisco Giants, Dynamic Pricing Software Hits a Home Run," CIO, June 29, 2011, at www.cio.com/article/685312.

One issue faced by managers long ago involved the facilitation of economic exchange. Whereas two individuals can negotiate face-to-face, coordination costs mount quickly as more people join. So managers had to devise a plan to reduce coordination costs and encourage more trade. They chose to create a social institution called a **market**.

A market exists when there is economic exchange; that is, multiple parties enter binding contracts. Countless markets exist in the world. The business world operates within these markets, and we need to examine (and understand) behavior in them. Surprisingly, given the number and diversity of markets, they all follow general principles. It is these principles we now focus on because knowing them is essential to understanding market behavior. We examine the behavior of individuals who enter contracts and on the aggregate effect that they create.

Market A group of firms and individuals that interact with each other to buy or sell a good.

THE DEMAND SIDE OF A MARKET

Every market consists of demanders and suppliers. A manager needs to know how potential customers value a product or service, and must estimate the quantity of goods demanded at various prices. One goal of managers is to maximize firm value. The ability to focus on profit requires a thorough knowledge of demand, especially the behavior of revenue as price changes. Total revenue is equal to the number of units sold (Q) multiplied by the price (P) at which they were sold ($TR = P \times Q$).

The association of price and quantity demanded often depends on many variables, some controlled by the manager and some not. Possible influences include income and tastes, prices of substitutes and complementary products, advertising dollars, product quality (as well as the quality of substitutes and complements), and governmental fiat. The behavior of quantity demanded relative to price is called a firm's **demand function** (holding other possible influences constant).

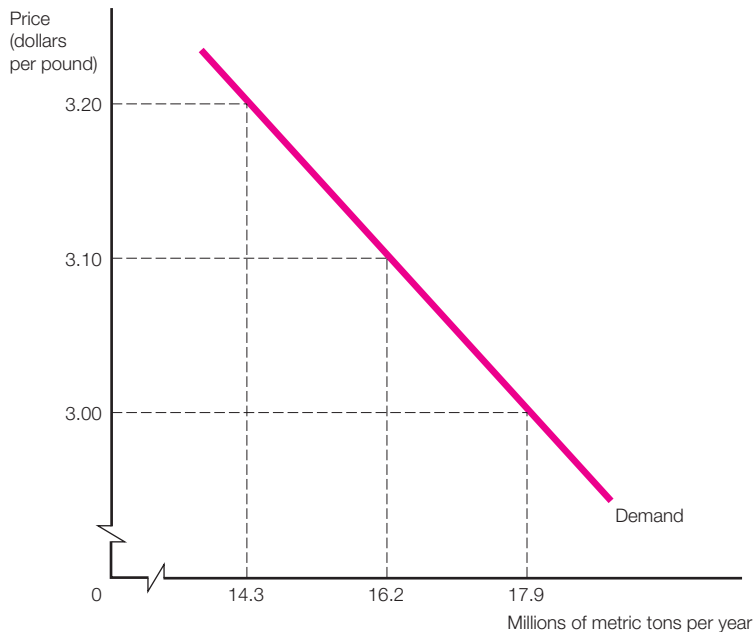
A demand curve shows managers how many units they sell at a given price. Consider Figure 1.1, which shows the demand curve for copper in the world market

Demand function Quantity demanded relative to price, holding other possible influences constant.

FIGURE 1.1

The Market Demand Curve for Copper, World Market

The market demand curve for copper shows the amount of copper that buyers would like to purchase at various prices.



STRATEGY SESSION: Supply and Demand in Action

On any day, view the commodity pages in the financial section of the print media or online, and you'll see the law of supply and demand at work.

For instance, in *The Economic Times* of India on May 11, 2011, the headline read "Pineapple Prices Fall on Over Supply." Reading on in the story revealed that pineapple prices had "crashed" (about a 40% decline) following a "glut" of pineapples arriving in the market. The number of loads arriving in Mumbai (Bombay) had jumped threefold. In addition, other summer fruit, such as mangos and watermelon, were starting to come to market, suppressing the demand for pineapple. A rightward shifting supply curve and a leftward shifting demand curve will lead to a decline in price—just as the text analysis tells us. Last year, pineapple prices rose 20%. This contributed to the 2011 supply shift as pineapple producers planted more in 2011.

On May 13, 2011, four headlines and several smaller stories appeared in *The Economic Times* of India: "Crude Oil Climbs Above \$100 on EU GDP Growth" (a rightward demand shift); "Speculation on Demand Pushed up Copper" (a rightward demand shift); "Brazil's Production Fall Lifts Sugar" (a leftward supply shift); "Buying Koreans Fuel US Meat Rally" (a rightward demand shift); and "Corn Rises in US on Rains" (a leftward supply shift).

U.S. beef prices have risen 13% in a year because of the added demand of Korean, Mexican, and Japanese consumers. While watching India and Thailand's sugar production, the world forgot about Brazil, where sugar production in one major region fell 69%. Brazil is the world's largest sugar producer. The headlines and the stories below them show that there are local markets (pineapples) and global markets (sugar, beef, copper, oil), and that one must be aware of supply-and-demand shifts both locally and internationally if one wants to predict prices.

But it's not just traditional consumption good commodities that exemplify the supply-demand model anymore. In an electronic world, more goods are now in active real-time supply-demand markets. For instance, the Boston Red Sox are normally a very strong baseball team. They should be: Their payroll is second to only the New York Yankees. But they started off the 2011 baseball season miserably, losing their first six games, all away from Boston. Their home games are usually sold out, but a secondary market exists in which people who have bought game tickets for speculative purposes or cannot attend a game for which they hold a ticket can offer these tickets for sale.

The first home game of the season, the season opener, is always a game where the demand for tickets is high. In addition, the first home game in 2011 was against the Red Sox's hated rival, the Yankees. However, ticket prices for the game were down about \$100 according to Ace Ticket, the largest Red Sox ticket broker. FanSnap.com had 5,000 tickets for the game listed for sale at an average price of \$253 on April 7, 2011, down about 23% from the April 1 average.

Two things likely happened here. First, demand shifted to the left as fickle fans are less enthusiastic about going to a game when the team is going bad (and the weather in Boston in early April is always iffy), and second, the supply is likely to shift rightward as ticket-holding fans are less enthusiastic about watching an underperforming team in the cold.

Sources: *The Economic Times*, May 11, 2011 and May 13, 2011; and "Boston Red Sox Prices Slump 40% as 0-6 Start Is Worst Since 1945," *Bloomberg News*, April 8, 2011, at www.bloomberg.com/news/2011-04-07/boston-red-sox-start-season-0-6-for-first-time-since-45-in-cleveland-loss.html.

in 2012. The figure shows that 16.2 million metric tons of copper are demanded annually if the price is \$3.10 per pound; 14.3 million metric tons if the price is \$3.20 per pound; and 17.9 million metric tons if the price is \$3.00 per pound. An important reason why copper has experienced recent growth in quantity demanded is its increasing use in emerging markets like China and India.

The demand curve in Figure 1.1 shows the global quantity of copper demanded at all prices. Any demand curve pertains to a particular period of time, and the shape and position of the demand curve can depend on the period length. The demand curve for copper slopes downward to the right. In mathematical terms, we say it has a *negative slope*; that is, the quantity of copper demanded increases as the price falls. This is true for most commodities: They almost always slope downward to the right. This makes sense; managers should expect price increases to result in lower sales.

Any demand curve is based on the assumption that other influences like tastes and incomes are held constant. Changes in any of these factors are likely to shift the position of a commodity's demand curve. So if consumers' tastes shift toward goods that use considerable copper or if consumers' incomes increase (and they thus buy more goods using copper), the demand curve for copper will shift to the right. In other words, holding the price of copper constant, more copper is demanded at any price. We will discuss this more fully in Chapter 2.

THE SUPPLY SIDE OF A MARKET

The supply side of a market is represented by a market supply curve that shows how many units of a commodity sellers will offer at any price. Figure 1.2 shows the supply curve for copper in the world market in 2012. According to the figure, 16.2 million metric tons of copper are supplied if the price of copper is \$3.10 per pound, 17.4 million tons if the price is \$3.20 per pound, and 14.9 million tons if the price is \$3.00 per pound.

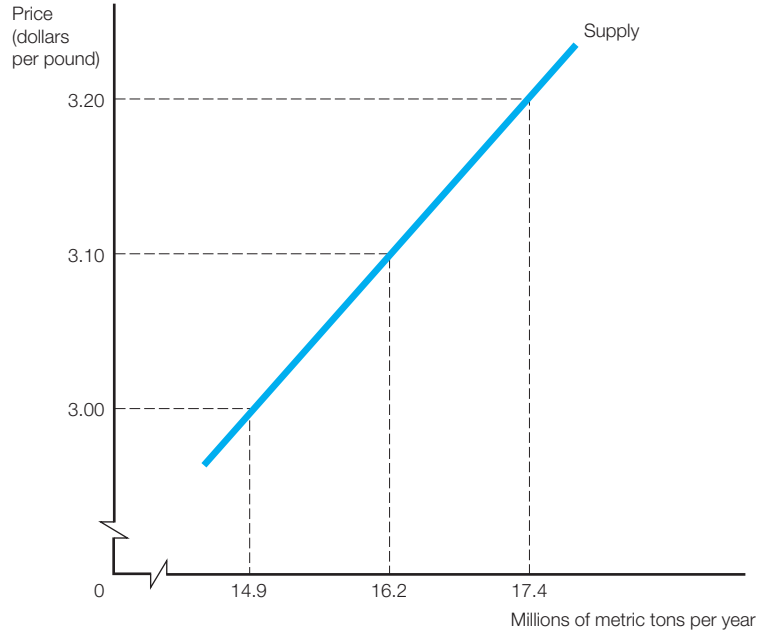
Note the supply curve slopes upward to the right. In mathematical terms, we say it has a *positive slope*; in other words, the quantity of copper supplied increases as the price rises. This seems plausible: Higher prices provide an incentive to suppliers to produce more copper to sell. Any supply curve is based on the assumption that production technology is held constant. If lower-cost production technology is developed, then managers will be willing to sell more units at any price. That is, technological change often causes a supply curve to shift to the right.

The supply curve for a product is affected by the cost of production inputs (labor, capital, and land). When costs of inputs decrease, managers realize lower production costs and are willing to supply a given amount at a lower price. So decreases in the cost of inputs cause supply curves to shift to the right. If input costs increase, managers are willing to supply a given amount only at a higher price (because their costs are higher). Hence the supply curve shifts to the left.

FIGURE 1.2

The Market Supply Curve for Copper, World Market

The market supply curve for copper shows the amount of copper that sellers would offer at various prices.

**EQUILIBRIUM PRICE**

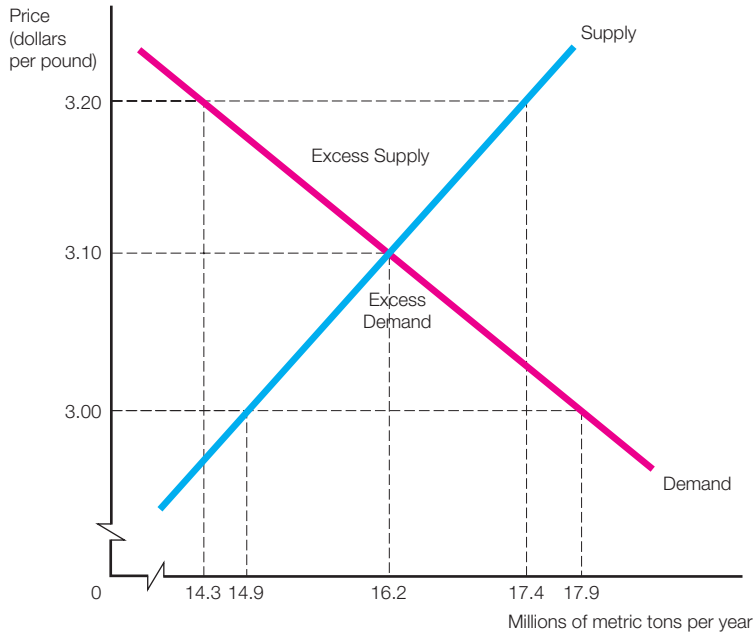
Economists represent markets as the interaction of demand and supply curves. To illustrate, consider the world copper market shown in Figure 1.3. We construct the figure by overlaying the demand curve (Figure 1.1) with the supply curve (Figure 1.2). Now we can determine market behavior at various prices. For example, if the price of copper is \$3.20 per pound, the demand curve indicates that 14.3 million metric tons of copper are demanded, while the supply curve indicates that 17.4 million metric tons are supplied. Therefore, if the market price is \$3.20 per pound, there is a mismatch between the quantity supplied and the quantity demanded. Specifically, as shown in Figure 1.3, there is excess supply of 3.1 million metric tons. Some producers will not be able to sell all their inventories at this price; they may be tempted to cut their prices to reduce these inventories. Hence a market price of \$3.20 per pound creates an unbalance in the market—there is too much supply. Because of this excess supply, producers will drop their prices, so \$3.20 is not a sustainable market price.

If the price is \$3.00 per pound, the demand curve indicates that 17.9 million metric tons are demanded, while the supply curve indicates that 14.9 million

FIGURE 1.3

Equilibrium Price of Copper, World Market

The equilibrium price is \$3.10 per pound, since quantity demanded equals the quantity supplied at this price.



metric tons are supplied. So a market price of \$3.00 also creates an unbalance in the market: There is not enough supply to satisfy demand. In fact, at this price consumers want to purchase an additional 3 million metric tons, but they can find no suppliers. When suppliers realize there is a shortage of copper they will increase their prices. Hence \$3.00 is not a sustainable market price.

So what is a sustainable market price? A price is sustainable when the quantity demanded at a price is equal to the quantity supplied at that price. The market is in balance because individuals who want to purchase the good can, and everyone who wants to sell the good can. We say the market is at **equilibrium**. In Figure 1.3, the price at which the quantity supplied equals the quantity demanded is \$3.10 per pound. This is also the point where the two curves intersect.

Equilibrium When the market is in balance because everyone who wants to purchase the good can and every seller who wants to sell the good can.

ACTUAL PRICE

Of course, price managers are interested in the actual price—the price that really prevails—not the equilibrium price. In general economists assume the actual price approximates the equilibrium price, which seems reasonable enough because the

STRATEGY SESSION: Using the Law of Supply and Demand to Invest in Stocks

Tax incentives designed to encourage homeowners and businesses in Italy and Germany are expiring, leading to a leftward shift in the demand curve for solar panels. At the same time, China's JA Solar Holdings Co. and Suntech (the two largest solar cell makers by capacity in the world) are greatly expanding their production capacity (as were other solar panel makers). This has led the supply curve to shift to the right. As discussed in this chapter, a leftward demand shift alone will cause prices to fall, and a rightward shift in supply alone will cause prices to fall. Put the two together, and prices really fall.

Investors watch for these market shifts. These price decreases do not bode well for firms making solar panels. Those with high production costs would be most hurt, but those with low production costs would still see a drop in profit margin. Savvy investors know that stock prices are generally positively correlated with profits, so investors sold solar panel companies' stocks short, with the intention of buying back shares on the cheap when solar panel firms' profits fell and hence stock prices fell.

First Solar Inc. of Tempe, Arizona, the world's largest producer of thin film solar panels, had 23% of their outstanding shares sell short. Q-Cells SE, a German company, had 54% of its shares sold short. Nine other large producers have large amounts of shares sold short (although none as much as First Solar and Q-Cells SE). The solar industries' stocks have a history of volatility, so much so that Shawn Kravetz of Esplanade Capital (a Boston hedge fund) and his colleagues have dubbed the industry: the "solarcoaster."

How does one make money by knowing about supply and demand? Watch for predicted price shifts and then buy or sell the product itself in the market, or buy or sell the stock in the market and cash out in the futures market. Sound easy? It actually takes nerve to be a speculator.

Source: "[BN] Shorts Sell 'Solarcoaster' as China Glut Sinks Panel Prices," *Bloomberg News*, June 19, 2011, at www.bloomberg.com/news/2011-06-19/short-sellers-hammer-solarcoaster-as-glut-of-chinese-panels-sinks-prices.html.

basic forces at work tend to push the actual price toward the equilibrium price. Therefore, if conditions remain fairly stable for a time, the actual price should move toward the equilibrium price.

To see this, consider the global market for copper, as described in Figure 1.3. What if the actual price of copper is \$3.20 per pound? As we have seen, this price will cause downward pressure on the price of copper. Suppose the price, responding to this pressure, falls to \$3.15 per pound. Comparing the quantity demanded with that supplied at \$3.15 per pound, we see there is still downward pressure on price because supply exceeds demand. The price, responding to this pressure, may fall to \$3.12 per pound; but comparing the quantity demanded with that supplied at this price, we find there is still downward pressure on price.

As long as the market price is greater than the equilibrium price, there is downward pressure on price. Similarly, as long as the actual price is less than the equilibrium price, there is upward pressure on price. Hence there is always a ten-

dency for the actual price to move toward the equilibrium price. The speed of this adjustment can vary. Sometimes it takes a long time for the actual price to approach the equilibrium price, and sometimes it happens quickly.

This price adjustment process is what Adam Smith called the market's **invisible hand**. No governmental agency is needed to induce producers to drop or increase their prices. They act more or less in unison and cause the market price to change.

Invisible hand When no governmental agency is needed to induce producers to drop or increase their prices.

WHAT IF THE DEMAND CURVE SHIFTS?

Any supply and demand diagram like Figure 1.3 is essentially a snapshot of the situation at a particular time. The results in Figure 1.3 are limited to a particular period because demand and supply curves are not static; they shift in reaction to changes in the environment. What happens to the equilibrium price of a good when its demand curve changes? This is important to know because managers need to anticipate and forecast price changes.

To illustrate the effects of a rightward shift of a demand curve, consider the copper industry in 2010. Housing starts were increasing (copper tubing is often used for water lines) in the emerging markets of China and India. As indicated in the right panel of Figure 1.4, managers should have expected that such a rightward shift of the demand curve would cause an increase in the price of copper from P

FIGURE 1.4

Effects of Leftward and Rightward Shifts of the Demand Curve on the Equilibrium Price of Copper

A leftward shift of the demand curve results in a decrease in the equilibrium price; a rightward shift results in an increase in the equilibrium price.



STRATEGY SESSION: Life During a Market Movement

When the market moves, planets can tremble. In early spring 2008, shifting demand and supply curves were impacting every country. The global food supply was in disequilibrium, and it appeared that the world was in a panic. During one week in early spring 2008, major governments worldwide used their sovereign powers to restrict trade in basic foods. Saudi Arabia cut import taxes on basic foodstuffs, India removed tariffs on edible oils while banning rice exports, and Vietnam cut its rice exports by 22%. Political unrest was beginning to erupt in countries as diverse as Egypt and Mexico over the rising cost of food. What is perhaps a mere inconvenience of paying \$4.25 for a gallon of milk in the United States is one of life or death to those who live on the fringes. For the 300 million Chinese who live in poverty, food accounts for 50% of household expenses.

Look at what caused this commotion. The figure on the next page shows the behavior of food prices prior to mid-2008.

You can clearly see the acceleration of prices across major food groups. This is a shift in the demand

curve. The *Financial Times* reported that the Philippine government paid \$700 per ton of rice—almost double the price the government paid in December 2007. The price of corn increased by 73% between April 2007 and April 2008 (from \$3.46 a bushel to \$6 a bushel). Wheat increased by 123% in this same period (from \$5.63 to \$12.57 a bushel). In China, the price of pork increased by 63%. In early 2012, a bushel of corn cost \$6.42 while a bushel of wheat was priced at \$6.48.

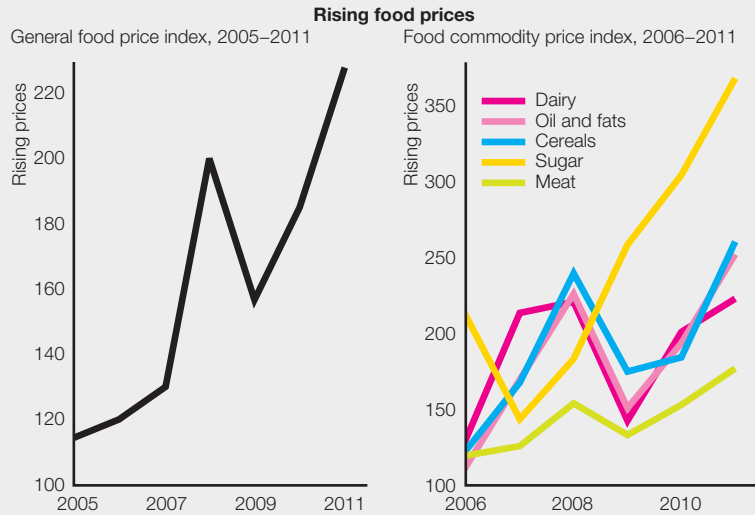
This rightward shift in the demand curve for food was attributed to several factors. One theory was that Thomas Malthus's mathematical doomsday machine was finally reaching fruition. The world population continues to expand, while agricultural acreage continues to shrink. Many governments in developing countries have focused efforts on economic development rather than agriculture. A UN report shows the annual growth in agricultural productivity slowed to 1% by 2002. A growing middle class in large developing countries like China and India consumes more food. As people increase their income, they generally eat more food. In China, consumption of meat has

to P_2 . In fact, the global price of copper in 2010 was roughly \$2.65 per pound. By 2012 this price had increased to \$3.82.

In mid-2009, we see a leftward shift in the demand curve for copper, as shown in the left panel of Figure 1.4. Because of slow economic growth in the United States and other countries, there was less demand for copper. This meant that the demand curve for copper shifted left, so there was less quantity demanded at any given price. Figure 1.4 shows a decrease in price from P to P_1 .

WHAT IF THE SUPPLY CURVE SHIFTS?

What happens to the equilibrium price of a product when its supply curve changes? For example, suppose that because of technological advances in cop-



doubled as personal income has increased. Finally, the increased use of food stock, like corn, for the production of ethanol fuel has taken such products out of the food market. Also, more individuals are leaving rural areas (farms) in developing countries and moving to urban areas. These trends are apt to

move the supply curve to the left and the demand curve to the right and put further upward price pressure on food.

Sources: "Countries Rush to Restrict Trade in Basic Foods," *Financial Times*, April 2, 2008, p. 1; "Food Prices Give Asian Nations a Wake-Up Call," *Financial Times*, April 3, 2008, p. 4.

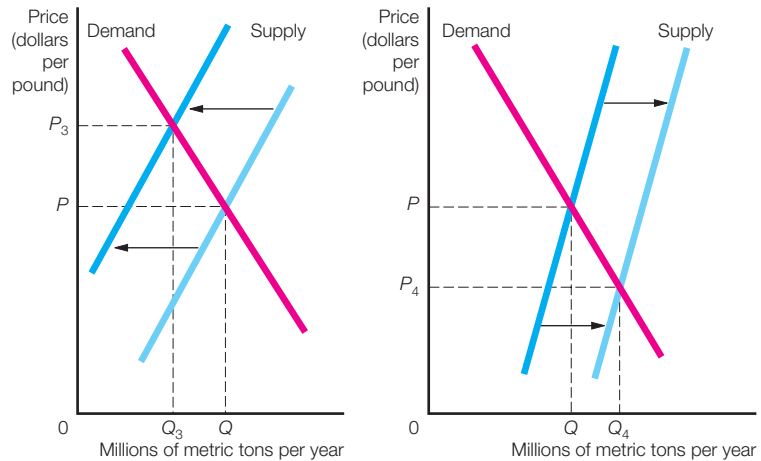
per production, large producers like Codelco of Chile can supply more copper at a given price than they used to. This will cause the supply curve to shift to the right, as shown in the right panel of Figure 1.5. How will this shift affect the equilibrium price? Clearly it will fall from P (where the original supply curve intersects the demand curve) to P_4 (where the new supply curve intersects the demand curve).

On the other hand, suppose there is a significant increase in the wage rates of copper workers. This increase will cause the supply curve to shift to the left, as shown in the left panel of Figure 1.5. This shift will cause the equilibrium price to increase from P (where the original supply curve intersects the demand curve) to P_3 (where the new supply curve intersects the demand curve).

FIGURE 1.5

Effects of Leftward and Rightward Shifts of the Supply Curve on the Equilibrium Price of Copper

A leftward shift of the supply curve results in an increase in the equilibrium price; a rightward shift results in a decrease in the equilibrium price.



SUMMARY

1. The main task of managers is to make decisions. We offer a guide for the managerial world; it is based on the behavioral economics of the managerial model. In contrast to microeconomics, which is largely descriptive, managerial economics is prescriptive. Courses in managerial economics provide fundamental analytical tools as well as play a major integrating role. Our decision framework describes behavior found at a wide range of organizations, from non-business organizations like government agencies to single-owner entrepreneurial efforts.

2. To apply economics to managerial behavior, we need a theory of the firm. The theory accepted by most managerial economists is that the owners want to maximize its value, defined as the present value of its expected future net cash flows (which for now are equated with profit). However, this maximization occurs subject to constraints because the firm has limited inputs, particularly in the very short run, and must comply with a variety of laws and contracts.

3. Managerial economists define profit somewhat differently from the way accountants do. When economists speak of profit, they mean profit over and

above what the owners' labor and capital employed in the business could earn elsewhere. To a considerable extent, the differences between the concepts of profit used by the accountant and the economist reflect the difference in their functions.

4. Three important reasons for the existence of profit are innovation, risk, and market power. Profit and loss are the mainspring of a free enterprise economy. They are signals showing where resources are needed and where they are too abundant. They are important incentives for innovation and risk taking. They are society's reward for efficiency.

5. Although managerial economists generally assume that owners want to maximize profit (and hence their value), a principal-agent problem arises if managers pursue their own interests, even though this decreases the profit of the owners. To address this problem, owners often give managers a financial stake in the future success of a firm.

6. Every market has a demand side and a supply side. The market demand curve shows the amount of a product buyers will purchase at various prices. The market supply curve shows the amount of a product producers are willing to sell at various prices. The equilibrium price is the price where the quantity demanded equals the quantity supplied. This price is also called the market-clearing price.

7. Both demand curves and supply curves can shift over time. This results in changes in a product's price. Rightward shifts in the demand curve (and leftward shifts in the supply curve) tend to increase price. Leftward shifts in the demand curve (and rightward shifts in the supply curve) tend to decrease price.

PROBLEMS

1. A book is to be written by Britney Spears. Batman Books agrees to pay Britney \$6 million for the rights to this not-yet-written memoir. According to one leading publisher, Batman Books could earn a profit of roughly \$1.2 million if it sold 625,000 copies in hardcover. On the other hand, if it sold 375,000 copies, managers would lose about \$1.3 million. Publishing executives stated that it was hard to sell more than 500,000 copies of a nonfiction hardcover book, and very exceptional to sell 1 million copies. Were Batman managers taking a substantial risk in publishing this book?
2. Some say that any self-respecting top manager joining a company does so with a front-end signing bonus. In many cases this bonus is in the seven figures. At the same time, the entering manager may be given a bonus guarantee. No matter what happens to firm profit, he or she gets at least a percentage of that bonus. Do long-term bonus guarantees help to solve the principal-agent problem, or do they exacerbate it? Why?

3. If the interest rate is 10%, what is the present value of the Monroe Corporation's profit in the next 10 years?

Number of Years in the Future	Profit (Millions of Dollars)
1	8
2	10
3	12
4	14
5	15
6	16
7	17
8	15
9	13
10	10

4. Managers at Du Pont de Nemours and Company expect a profit of \$2.9 billion in 2012. Does this mean that Du Pont's expected economic profit will equal \$2.9 billion? Why or why not?
5. William Howe must decide whether to start a business renting beach umbrellas at an ocean resort during June, July, and August of next summer. He believes he can rent each umbrella to vacationers at \$5 a day, and he intends to lease 50 umbrellas for the three-month period for \$3,000. To operate this business, he does not have to hire anyone (but himself), and he has no expenses other than the leasing costs and a fee of \$3,000 per month to rent the business location. Howe is a college student, and if he did not operate this business, he could earn \$4,000 for the three-month period doing construction work.
- If there are 80 days during the summer when beach umbrellas are demanded and Howe rents all 50 of his umbrellas on each of these days, what will be his accounting profit for the summer?
 - What will be his economic profit for the summer?
6. On March 3, 2008, a revival of *Gypsy*, the Stephen Sondheim musical, opened at the St. James Theater in New York. Ticket prices ranged from \$117 to \$42 per seat. The show's weekly gross revenues, operating costs, and profit were estimated as follows, depending on whether the average ticket price was \$75 or \$65

	Average Price of \$75	Average Price of \$65
Gross revenues	\$765,000	\$680,000
Operating costs	600,000	600,000
Profit	165,000	80,000

- a. With a cast of 71 people, a 30-piece orchestra, and more than 500 costumes, *Gypsy* cost more than \$10 million to stage. This investment was in addition to the operating costs (such as salaries and theater rent). How many weeks would it take before the investors got their money back, according to these estimates, if the average price was \$65? If it was \$75?
 - b. George Wachtel, director of research for the League of American Theaters and Producers, has said that about one in three shows opening on Broadway in recent years has at least broken even. Were the investors in *Gypsy* taking a substantial risk?
 - c. According to one Broadway producer, "Broadway isn't where you make the money any more. It's where you establish the project so you can make the money. When you mount a show now, you really have to think about where it's going to play later." If so, should the profit figures here be interpreted with caution?
 - d. If the investors in this revival of *Gypsy* make a profit, will this profit be, at least in part, a reward for bearing risk?
7. If the demand curve for wheat in the United States is

$$P = 12.4 - Q_D$$

where P is the farm price of wheat (in dollars per bushel) and Q_D is the quantity of wheat demanded (in billions of bushels), and the supply curve for wheat in the United States is

$$P = -2.6 + 2Q_S$$

where Q_S is the quantity of wheat supplied (in billions of bushels), what is the equilibrium price of wheat? What is the equilibrium quantity of wheat sold? Must the actual price equal the equilibrium price? Why or why not?

8. The lumber industry was hit hard by the downturn in housing starts in 2010 and 2011. Prices plunged from \$290 per thousand board feet to less than \$200 per thousand board feet. Many observers believed this price decrease was caused by the slowing of new home construction because of the glut of unsold homes on the market. Was this price decrease caused by a shift in the supply or demand curve?
9. From November 2010 to March 2011 the price of gold increased from \$1,200 per ounce to over \$1,800 per ounce. Newspaper articles during this period said there was little increased demand from the jewelry industry but significantly more demand from investors who were purchasing gold because of the falling dollar.
 - a. Was this price increase due to a shift in the demand curve for gold, a shift in the supply curve for gold, or both?
 - b. Did this price increase affect the supply curve for gold jewelry? If so, how?

EXCEL EXERCISE: DEMAND, SUPPLY, AND MARKET EQUILIBRIUM

Suppose you were given the following information about market demand:

Price (P)	Quantity Demanded (Q_D)
14	0
13	1
12	2
11	3
10	4
9	5
8	6
7	7
6	8
5	9
4	10
3	11
2	12
1	13
0	14

and the following information about market supply:

Price (P)	Quantity Supplied (Q_S)
2	0
3	1
4	2
5	3
6	4
7	5
8	6
9	7
10	8
11	9
12	10
13	11
14	12

You could predict market equilibrium in several ways.

Suppose you noticed that every number in the demand table corresponded to numbers generated by the formula

$$Q_D = 14 - P$$

Your statistics course will teach you how to fit an equation to data, but you may have encountered an SAT, GMAT, or GRE question that gave you a column of numbers and then asked you to continue the column for another one or two entries. If you answered that problem correctly, you implicitly solved for the above equation.

Suppose you noted that every number in the supply table corresponded to numbers generated by the formula

$$Q_S = -2 + P$$

You know from the text that in equilibrium, $Q_D = Q_S$, and so setting the two equal and solving for P gives

$$Q_D = 14 - P = -2 + P = Q_S$$

or

$$2P = 16$$

or

$$P = 8$$

Substituting $P = 8$ into the demand equation gives

$$Q_D = 14 - 8 = 6$$

Substituting $P = 8$ into the supply relationship gives

$$Q_S = -2 + 8 = 6$$

Thus market equilibrium is $P = 8$ and $Q = 6$.

But suppose that you couldn't estimate the demand and supply equations. Can you use a spreadsheet to calculate the equilibrium? Yes. Here's how.

Open your spreadsheet and put 14 in cell A1, then 13 in cell A2, and 12 in cell A3, and so on. Thus, you will have entered each price sequentially from 14 to 0.

Then put 0 in cell B1, 1 in cell B2, 2 in cell B3, and so on. Thus, you will have entered all the quantities demanded at their corresponding price.

Then put 12 in cell C1, 11 in cell C2, 10 in cell C3, down to 0 in cell C13. Thus, you will have entered all quantities supplied at their corresponding price.

Then in cell D1 enter the formula $= B1 - C1$ (Note: You must have the $=$ sign, and don't skip spaces) and in cell D2 enter the formula $= B2 - C2$, and so on. You don't actually have to enter the formulas multiple times. Just click on the cell D1 in the lower righthand corner, drag with your mouse down to cell D13, and the correct formulas will be transferred to cells D2 to D13.

The numbers that will appear in cells D1 to D13 will be the quantity demanded (Q_D) at the respective price minus the quantity supplied (Q_S) at the respective

price (called the *excess demand*). As there is no excess demand at equilibrium, that is $Q_D = Q_S$, you look for the D column cell with a 0. If you've entered everything correctly, it'll be cell D7. Reading leftward across the seventh row, you'll see that $P = 8$ and $Q_D = Q_S = 6$.

The nice thing about the spreadsheet is that you can clearly see the amount of excess demand at each price above the market equilibrium, for example -10 at $P = 12$, and so you can imagine the pressure on the suppliers to lower the price to get rid of that 10 of inventory.

When the excess demand is negative (above the market equilibrium price), we usually multiply it through by -1 to make it positive and call it *excess supply*. Doing so enables us to see by how much Q_S exceeds Q_D when the market price is too high. When there is positive excess demand, the positive number shows the motivation of unsatisfied demanders with higher reservation prices to up their bids for the product (and the motivation for sellers to raise their prices).

Throughout the book, we will present other applications of using the spreadsheet to solve the economic problems facing managers.

PART 2

THE NATURE OF MARKETS

CHAPTER 2

LEARNING OBJECTIVES

The Market Demand Curve

Industry and Firm Demand Functions

The Own-Price Elasticity of Demand

Point and Arc Elasticities

Using the Demand Function to Calculate the Price Elasticity of Demand

The Effect of Price Elasticity on the Firm's Revenue

Funding Public Transit

Determinants of the Own-Price Elasticity of Demand

The Strategic Use of the Price Elasticity of Demand

Total Revenue, Marginal Revenue, and Price Elasticity

The Income Elasticity of Demand

Cross-Price Elasticities of Demand

The Advertising Elasticity of Demand

The Constant-Elasticity and Unitary Elastic Demand Function

Summary

Problems

DEMAND THEORY

An important determinant of profit is the nature of the demand for a firm's goods or services. It is imperative that managers understand this multidimensional concept if they are to positively influence firm performance. Good managers learn to understand the nature of demand for products and effectively manage it. Effective management requires more knowledge than understanding the directional impact on sales for a given price change. Many other factors besides price affect consumer demand. Some of these factors are controlled by managers, such as advertising, product quality, and distribution. Other factors, like the number of substitute goods, the prices of rival products, and the advertising of rivals, are part of the competitive dynamics of the product space. Finally, a few factors, like the state of the economy or the level of disposable consumer income, are macroeconomic and are not influenced by individual managers. Though these factors are outside their control, managers still need to predict how their changes affect demand.

This chapter explains how managers can more precisely predict changes in various environmental factors and quantify their impact on product demand. The nature of product demand is that it is a process—and as such is dynamic. Because many factors influence product demand, managers need to understand how changes in these factors affect demand. Knowing the sensitivity of demand to changes in environmental factors lets a manager effectively respond to these changes. The sensitivity of one factor to another is called **elasticity**. Elasticity measures the percentage change in one factor given a small (marginal) percentage

Elasticity Elasticity measures the percentage change in one factor given a small (marginal) percentage change in another factor.

change in another factor. The concept of elasticity is widespread in the business world. For example, elasticity is the basis for both a firm's operating and financial leverage. It is also used by managers to determine a product's most efficient mix of inputs.

THE MARKET DEMAND CURVE

One way to show how sales of a product are affected by its price is with a **market demand schedule**, which is a table showing the total quantity of the good purchased at each price. For example, suppose the market demand schedule for tablet computers in 2012 is as shown in Table 2.1. According to this table, 0.5 million tablet computers are demanded per year if the price is \$200 per computer; 300,000 are demanded if the price is \$300; and so on. Another way of presenting the data in Table 2.1 is with a **market demand curve**, which is a plot of the market demand schedule on a graph. The vertical axis of the graph measures the price per unit of the good, and the horizontal axis measures the quantity of the good demanded per unit of time. Figure 2.1 shows the market demand curve for tablet computers in 2012, based on the figures in Table 2.1.

In the previous chapter we introduced the concept of a market demand curve. Now we examine one in more detail. Note three things about Figure 2.1. First, the market demand curve shows the total quantity of tablet computers demanded at each price, not the quantity demanded from a particular firm. We discuss the demand for a particular firm's product later. Second, the market demand curve for tablets slopes downward to the right. That is, the quantity of tablets demanded increases as the price falls. As we pointed out in the last chapter, this is true for most products or services. Third, the market demand curve in Figure 2.1 pertains to a specified period: 2012. As you recall from the last chapter, any demand curve pertains to some particular time, and its shape and position depend on the length

Market demand schedule Table showing the total quantity of the good purchased at each price.

Market demand curve The plot of the market demand schedule on a graph.

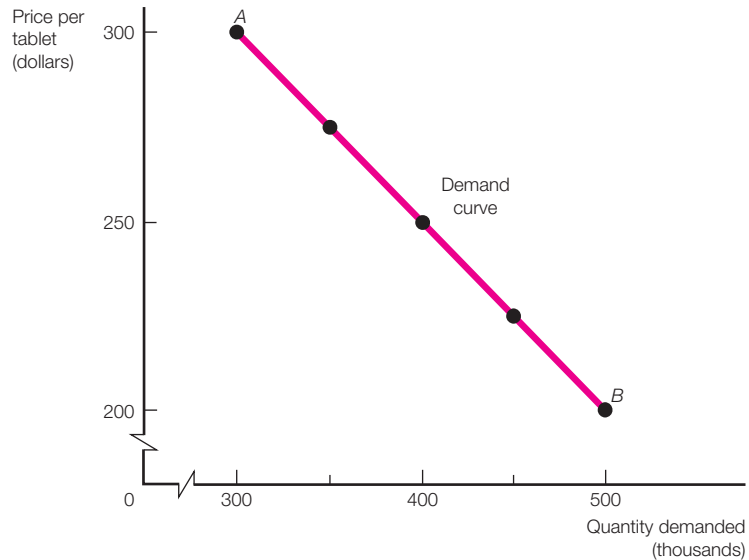
TABLE 2.1

Market Demand Schedule for Tablets, 2012

Price per Tablet (Dollars)	Quantity Demanded (Thousands)
300	300
275	350
250	400
225	450
200	500

FIGURE 2.1**Demand Curve for Tablets**

This demand curve is a graphical representation of the figures in Table 2.1.



and other characteristics of this period. If we were to estimate the market demand curve for tablets for the first week in 2012, it would probably differ from the one in Figure 2.1. This difference arises partly because consumers adapt their purchases more fully to changes in the price of tablets in a year relative to a week.

In addition to the length of time, what factors determine the position and shape of a market demand curve? One important factor is the tastes of consumers. If consumers show an increasing preference for a product, the demand curve shifts to the right; that is, at each price consumers want to buy more than they did previously. Alternatively, for each quantity consumers are willing to pay a higher price. On the other hand, if consumers show a decreasing preference for a product, the demand curve shifts to the left because at each price consumers buy less than previously. Alternatively, for each quantity consumers are willing to pay only a lower price. For example, as shown in Figure 2.2, if people find that tablets are helpful and begin to use them more and give them in larger numbers to their children and others, the demand curve may shift to the right. The greater the shift in preferences, the farther the demand curve shifts.

Another factor that influences the position and shape of a product's market demand curve is the level of consumer incomes. For some products the demand curve shifts to the right if per capita income increases, whereas for other products

STRATEGY SESSION: The Customer Is Always Right—Wrong!

Some customers are viewed as angels. They purchase the big-ticket, big-markup items, want them right when they come out, and want each one of them. And some customers are viewed as devils. They wait for loss leader sales (items sold at a loss designed to entice consumers to the store, where the store hopes they will buy many other items—the devils only buy the loss leaders); buy items, return them, and then rebuy them as previously owned items at a discount; buy the most discounted items and then resell them (eliminating their customers from the store's pool); scour the Internet, circulars, and newspapers for the store's competitors' low prices and then make the store honor the competitors' prices because of the "we will not be undersold" pledge of the store; send in for rebates; and so on.

Big box electronics store Best Buy has had enough of the devils' tactics. Best Buy estimates that 20% of its store visits are by devils. And they'd like the devils to get out of their stores. They want to fire some of their customers! On the other hand, they have identified the true angels—the 20% of their customers that generate the bulk of their profits.

How do you get rid of the devils? You can't identify them and then not let them in your store. That would violate antidiscrimination laws. But you can eliminate the programs that draw them to your stores (and make sure they are not the same programs that draw the angels). For instance, stop direct mailing to the customers identified (by their past purchases) as devils; charge customers a restocking fee of 15% of the purchase price for returned items; prohibit reselling returned items on the Internet or at another store rather than at the original store; and break all ties with Internet sites (FatWallet.com, SlickDeals.net, TechBargains.com) that tipped the devils off to Best Buy bargains and buying strategies that Best Buy regarded as having a negative impact on profits. The financial services sector has solved this problem by catering to their angels (free checking for maintaining a certain balance) and penalizing the devils (transaction fees for ATM use, fees to deal with a teller, check fees, and so forth).

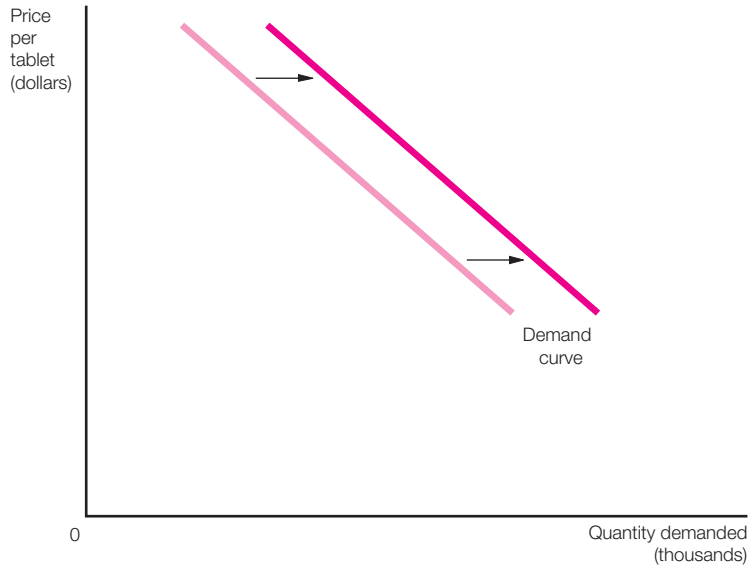
Source: "Analyzing Customers, Best Buy Decides Not All Are Welcome," *The Wall Street Journal*, November 8, 2004, p. A-1.

it shifts to the left if per capita income rises. In the case of tablets, we expect that an increase in per capita income will shift the demand curve to the right, as shown in Figure 2.3. Still another factor that influences the position and shape of a product's market demand curve is the level of other prices. For example, we expect the quantity of tablets demanded to increase if the price of applications falls drastically.

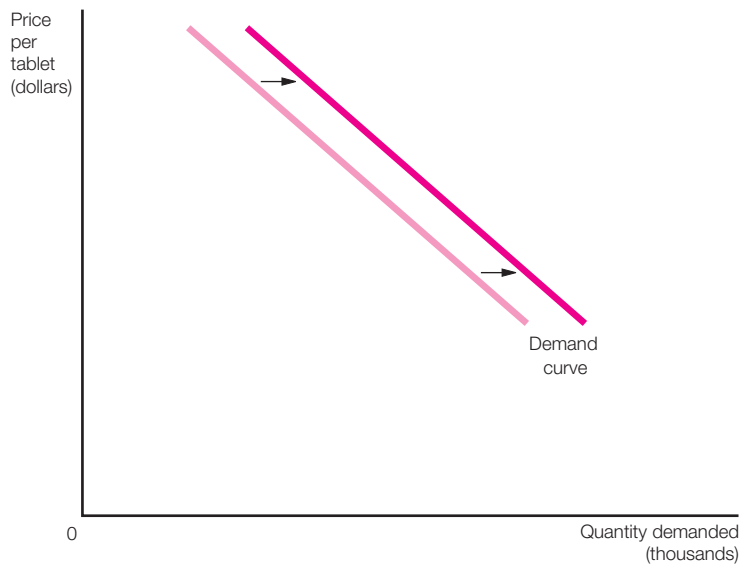
Finally, the position and shape of a product's market demand curve are affected by the size of the population in the relevant market. If the number of consumers increases, we expect that, if all other factors are held equal, the quantity of tablets demanded will increase. Of course, the population generally changes slowly, so this factor often has little effect in the short run.

FIGURE 2.2**Effect of an Increased Preference on the Market Demand Curve for Tablets**

The demand curve for tablets shifts to the right.

**FIGURE 2.3****Effect of an Increase in Per Capita Income on the Market Demand Curve for Tablets**

The demand curve shifts to the right.



INDUSTRY AND FIRM DEMAND FUNCTIONS

Building on the results of the previous section, we define the **market demand function** for a product as the relationship between the quantity demanded and the various factors that influence this quantity. Put generally, this market demand function is written as

$$\text{Quantity demanded of good } X = Q = f(\text{price of } X, \text{ incomes of consumers, tastes of consumers, prices of other goods, population, advertising expenditures, and so forth})$$

To be useful for analytical and forecasting purposes, this equation must be more specific. For example, if good X is tablet computers, the market demand function might be

$$Q = b_1P + b_2I + b_3S + b_4A \quad (2.1)$$

where in a particular year, Q equals the number of tablets demanded, P is the average price of tablets, I is per capita disposable income, S is the average price of applications for tablets and A is the amount spent on advertising by tablet producers. Equation (2.1) assumes that the relationship is linear. (Also, we assume the population in the relevant market is essentially constant.)

Going a step further, it is generally necessary for managers to obtain numerical estimates of the values of the b 's in equation (2.1). Employing various statistical techniques, managers estimate these **parameters** of the demand function to increase their knowledge of the demand for their product. To illustrate the results we might obtain, we find that

$$Q = -2,000P + 70I - 375S + 0.0001A \quad (2.2)$$

According to equation (2.2), a \$1 increase in the price of a tablet computer decreases the quantity demanded by 2,000 units per year; a \$1 increase in per capita disposable income results in a 70 unit increase in quantity demanded per year; a \$1 increase in the average price of applications reduces the quantity demanded by 375 units per year; and a \$1 increase in advertising raises the quantity demanded by 0.0001 units per year.

It is important to understand the relationship between the market demand function and the demand curve. The market demand curve shows the relationship between Q and P when all other relevant variables are held constant. For example, suppose we want to know the relationship between quantity demanded and price if per capita disposable income is \$13,000, the average price of applications is \$40,

Market demand function The relationship between the quantity demanded and the various factors that influence this quantity.

Parameters Constant or variable terms used in the function that help managers determine the specific form of the function but not its general nature.

and advertising expenditure is \$50 million. Because $I = 13,000$, $S = 40$, and $A = 50,000,000$, equation (2.2) becomes

$$Q = -2,000P + 70(13,000) - 375(40) + 0.0001(50,000,000) \quad (2.3)$$

Or
$$Q = 900,000 - 2,000P \quad (2.4)$$

Solving this equation for P , we obtain

$$P = 450 - 0.0005Q$$

which is graphed Figure 2.1. This is the demand curve for tablets, given that I , S , and A are held constant at the stipulated levels.

Given the market demand function, managers can better understand how changes in variables can shift the demand curve. For example, how much of a shift will occur in the demand curve if the average price of applications falls from \$40 to \$20? Inserting 20 (rather than 40) for S in equation (2.3), we find that

$$Q = 907,500 - 2,000P \quad (2.5)$$

Solving this equation for P , we obtain

$$P = 453.75 - 0.0005Q \quad (2.6)$$

STRATEGY SESSION: Servicing Demand in Emerging Markets

Direct selling exists in the United States, as is illustrated by companies such as Herbalife, Avon, and Tupperware, and in other developed markets. Both Herbalife and Avon have recently outperformed analyst estimates of earnings. In general, however, these businesses are the Rodney Dangerfields of business—that is, they get no respect from their business colleagues.

But in less-developed parts of the world where small-scale shopping is performed, such as in very small stores, on sidewalks, and on street corners with no agglomeration commerce or one-stop shopping possibilities, the opportunities of direct marketing seem to be ready to proliferate. The CEO of Tupperware, Rick Goings, reports that women around the

world are signing up “in droves” to be sales representatives, that is, holders of Tupperware parties. It’s hard to advertise broadly in less-developed places. In environments where limited experience with new products is common due to low income and chance for exposure levels, customers seem to value the added personal touch and the testimonials.

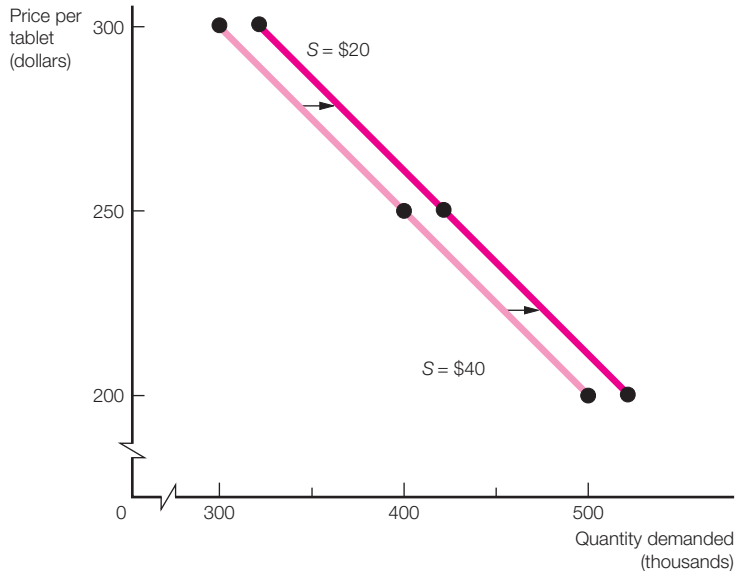
Although the megamall and new products may seem remote possibilities in the less-developed world, word of mouth and the ability to see and handle the product in a personal setting may be the future of sales in such areas.

Source: Michelle Fox, “The Best Way to Sell Goods in Emerging Markets?” at www.cnn.com, at www.cnn.com/id/42876221.

FIGURE 2.4

Demand Curve for Tablets

If the price of applications falls from \$40 to \$20, the demand curve shifts to the right by 7,500 units.



which is graphed (together with the demand curve based on $S = 40$) in Figure 2.4. Clearly the demand curve has shifted to the right: The quantity demanded is 7,500 more than when $S = 40$ (if P is held constant).

Managers are often more interested in the demand curves of their own brands rather than the market demand curve. We can derive these demand functions in a similar manner to predict the sales of an individual producer of tablet computers. In such an equation, the quantity demanded of the firm's product is still inversely related to its own price but is directly related to prices charged by its competitors. So if Dell increases its tablet prices, sales of Apple iPad 2 tablets will increase (all other factors being equal). It is important to distinguish between industry and firm demand functions because they are quite different. Although both are important for managers to understand, they are generally used for different purposes. Looking at the demand functions for individual firms is important in understanding the competitive dynamics of the market. Market demand curves tell managers more about substitute goods outside their product market and the general effects of macroeconomic factors like changes in disposable income on industry sales.

THE OWN-PRICE ELASTICITY OF DEMAND

The elasticity of a function is defined as the percentage change in the dependent variable in response to a 1% change in the independent variable. In $y = ax$, y is the dependent variable because by specifying x , we can determine y . A market demand curve is a function in which quantity demanded is dependent on a product's price. Market demand curves vary with regard to the sensitivity of quantity demanded to price. For some goods, a small price change results in a big change in quantity demanded; for other goods, a big price change results in a small change in quantity demanded. To indicate how sensitive quantity demanded is to price changes, economists use a measure called the **own-price elasticity of demand**. The word *own* is used to convey the idea that managers generally measure the price elasticity of demand for a product or service produced by their firm. More commonly, own-price elasticity of demand is simply referred to as the price elasticity of demand. The price elasticity of demand is defined as the percentage change in quantity demanded resulting from a 1% change in price. More precisely, it equals

$$\eta = \left(\frac{P}{Q}\right) \frac{\Delta Q}{\Delta P} \quad (2.7)$$

$\Delta Q/\Delta P$ (horizontal change/vertical change) is the inverse of a line's slope. Because linear demand curves are downward-sloping, $\Delta Q/\Delta P$ is negative; hence the price elasticity of demand is expressed as a negative number. Suppose a 1% reduction in the price of Apple tablet computers results in a 1.3% increase in U.S. sales. If so, the price elasticity of demand for Apple tablets is -1.3 . The price elasticity of demand generally changes as price varies along the demand curve. For instance, the price elasticity of demand may be higher in absolute value when the price of tablets is relatively high than when it is low. Similarly, the price elasticity of demand varies from market to market. India probably has a different price elasticity of demand for tablets than that of the United States.

We can classify the price elasticity of demand as falling into one of three possible "buckets." When a 1% change in price leads to a more than 1% change in quantity demanded, we say demand is **elastic**. When a 1% change in price leads to a less than 1% change in quantity demanded, we say demand is **inelastic**. And when a 1% change in price leads to a 1% change in quantity demanded, we say demand is **unitary elastic**. Because the price elasticity of demand is always negative for linear demand, we express this information as follows: When demand is elastic, $\eta < -1$; when demand is inelastic, $\eta > -1$; and when demand is unitary elastic, $\eta = -1$.

The price elasticity of demand for a product must lie between zero and negative infinity. If the price elasticity is zero, the demand curve is a vertical line; that is, the quantity demanded is unaffected by price. If the price elasticity is negative

Own-price elasticity of demand

More simply referred to as price elasticity of demand, this is the concept managers use to measure their own percentage change in quantity demanded resulting from a 1% change in their own price.

Elastic Elastic is used to describe demand when a 1% change in price leads to a more than 1% change in quantity demanded.

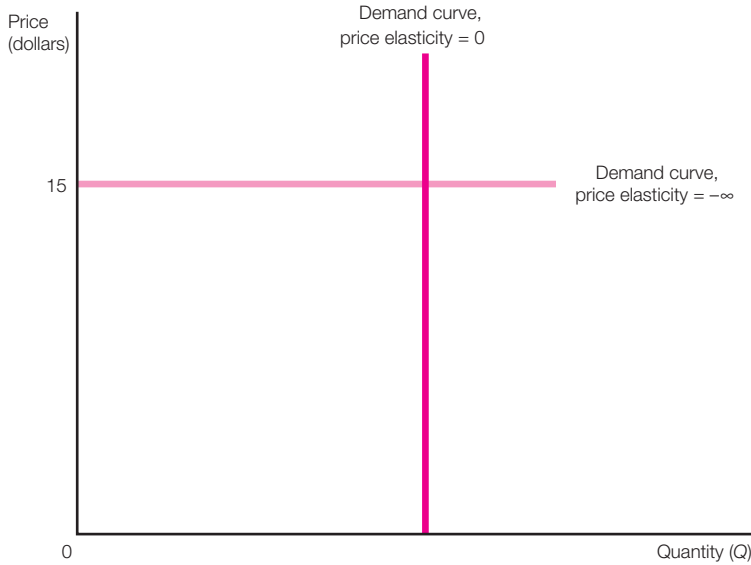
Inelastic Inelastic is used to describe demand when a 1% change in price leads to a less than 1% change in quantity demanded.

Unitary elastic Unitary elastic is used to describe demand when a 1% change in price leads to a 1% change in quantity demanded.

FIGURE 2.5

Demand Curves with Zero and Infinite Price Elasticities of Demand

The demand curve is a vertical line if the price elasticity is zero and a horizontal line if it is negative infinity.



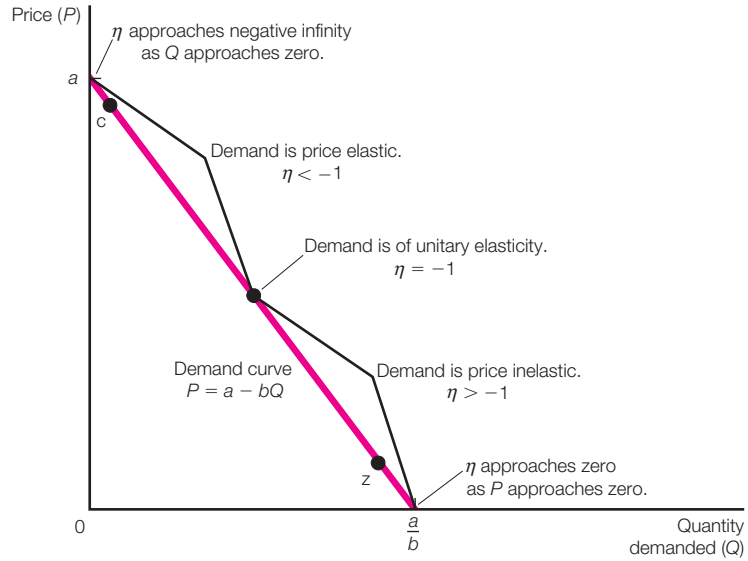
infinity, the demand curve is a horizontal line; that is, an unlimited amount can be sold at a particular price (\$15 in Figure 2.5), but nothing is sold if the price is raised even slightly. Figure 2.5 shows these two limiting cases.

We know that $\Delta Q / \Delta P$ is constant along a linear function. However, the price elasticity of demand even for linear demand is not constant because the ratio of price to quantity (P/Q) varies as we move along the demand curve. This is illustrated in Figure 2.6. Point c represents a point on the demand curve. Price is very high, and quantity demanded is very low. Hence P/Q is a large positive number, so elasticity must be low ($\eta < -1$, and demand is elastic). At point z the opposite is true. Price is very low, and quantity demanded is very high; hence P/Q is less than 1. So elasticity is high ($\eta > -1$, and demand is inelastic). Hence for linear demand curves, when P is high, the price elasticity of demand is large (in absolute value terms). As we move down the demand curve, P is decreasing and Q is increasing. This causes the price elasticity to monotonically increase. As we approach the horizontal axis, by definition, P is low and Q is high, so the demand is inelastic. Because the price elasticity is high (above -1) when P is low, and low (below -1) when P is high, at some point on any linear demand curve, the price elasticity must be equal to -1 (unitary elasticity).

FIGURE 2.6

Values of the Price Elasticity of Demand at Various Points along a Linear Demand Curve

The price elasticity increases in absolute value as price rises, approaching negative infinity as quantity approaches zero.



An alternative way of seeing how the price elasticity changes as we move along the demand curve is the following. If

$$P = a - bQ$$

where a is the intercept of the demand curve on the price axis and b is the slope (in absolute terms) of the demand curve, it follows that

$$Q = \frac{a}{b} - \frac{1}{b}P$$

Therefore, the price elasticity of demand is

$$\left(\frac{\Delta Q}{\Delta P}\right)\left(\frac{P}{Q}\right) = \left(\frac{-1}{b}\right)\frac{a - bQ}{Q}$$

Clearly, if the demand curve is linear, the price elasticity approaches zero as $P (= a - bQ)$ gets very small and approaches negative infinity as Q gets very small.

So for most demand curves, the elasticity of demand varies with price. That is, the magnitude of the sales response to price changes does not remain constant. This makes it more difficult for managers to estimate the effects of price changes on sales levels.

POINT AND ARC ELASTICITIES

If we have a market demand schedule showing the quantity of a product demanded at various prices, how can we estimate the price elasticity of market demand? Let ΔP be a change in the price of the good and ΔQ the resulting change in its quantity demanded. If ΔP is very small, we can compute the point elasticity of demand

$$\eta = \frac{\Delta Q}{Q} \div \frac{\Delta P}{P} \quad (2.8)$$

For instance, consider Table 2.2, where data are given for very small increments in the price of a commodity. If we want to estimate the price elasticity of demand when the price is between \$0.9995 and \$1, we obtain

$$\eta = \frac{20,002 - 20,000}{20,000} \div \frac{99.95 - 100}{100} = -0.2$$

Note that we used \$1 as P and 20,000 as Q . We could have used \$0.9995 as P and 20,002 as Q , but it would have made no real difference to the answer.

But if we have data concerning only large changes in price (that is, ΔP and ΔQ are large), the answer may vary considerably depending on which values of P and Q are used in equation (2.8). Consider the example in Table 2.3. Suppose we want to estimate the price elasticity of demand in the price range between \$4 and \$5. Then, depending on which values of P and Q are used, the answer is

$$\eta = \frac{40 - 3}{3} \div \frac{4 - 5}{5} = -61.67$$

TABLE 2.2

Quantity Demanded at Various Prices (Small Increments in Price)

Price (Cents per Unit of Commodity)	Quantity Demanded per Unit of Time (Units of Commodity)
99.95	20,002
100.00	20,000
100.05	19,998

TABLE 2.3

Quantity Demanded at Various Prices (Large Increments in Price)

Price (Dollars per Unit of Commodity)	Quantity Demanded per Unit of Time (Units of Commodity)
3	50
4	40
5	3

or

$$\eta = \frac{3 - 40}{40} \div \frac{5 - 4}{4} = -3.70$$

The difference between these two results is very large. To avoid this difficulty, it is advisable to compute the arc elasticity of demand, which uses the average values of P and Q

$$\begin{aligned}\eta &= \frac{\Delta Q}{(Q_1 + Q_2)/2} \div \frac{\Delta P}{(P_1 + P_2)/2} \\ \eta &= \frac{\Delta Q(P_1 + P_2)}{\Delta P(Q_1 + Q_2)}\end{aligned}\quad (2.9)$$

where P_1 and Q_1 are the first values of price and quantity demanded, and P_2 and Q_2 are the second set. Therefore, in Table 2.3,

$$\eta = \frac{40 - 3}{(40 + 3)/2} \div \frac{4 - 5}{(4 + 5)/2} = -7.74$$

USING THE DEMAND FUNCTION TO CALCULATE THE PRICE ELASTICITY OF DEMAND

Managers often estimate a demand function for their products. In equation (2.2) we provided the following hypothetical demand function for tablet computers

$$Q = -2,000P + 70I - 375S + 0.0001A$$

Given such a demand function, how can we calculate the price elasticity of demand?

The first step is to specify the point on the demand curve at which price elasticity is to be measured. Assuming per capita disposable income (I) is \$13,000, the

average price of applications is \$40, and advertising expenditure is \$50,000,000, we know from equation (2.4) that the relationship between quantity demanded and price is

$$Q = 900,000 - 2,000P \quad (2.10)$$

Suppose we want to measure the price elasticity of demand when price equals \$300. At this point on the demand curve (point A in Figure 2.1),

$$Q = 900,000 - 2,000(300) = 300,000$$

We can express $\Delta Q/\Delta P$ as the inverse of the slope. Hence in our example,

$$\Delta Q/\Delta P = -2,000 = -1/0.0005$$

According to equation (2.7), to obtain the price elasticity of demand, we must multiply $\Delta Q/\Delta P$ by P/Q . Performing this multiplication, we get

$$-2,000(300/300,000) = -2$$

which means the price elasticity of demand equals -2 .

As a further illustration, let us calculate the price elasticity of demand when price equals \$200 rather than \$300. At this point on the demand curve (point B in Figure 2.1),

$$Q = 900,000 - 2,000(200) = 500,000$$

Because $\Delta Q/\Delta P = -2,000$,

$$\eta = (\Delta Q/\Delta P)(P/Q) = -2,000(200/500,000) = -0.8$$

Therefore, the price elasticity of demand equals -0.8 .

THE EFFECT OF PRICE ELASTICITY ON THE FIRM'S REVENUE

As we have seen, estimating the price elasticity of demand helps managers predict how a given price change will affect sales. We can also use price elasticities to determine how a price change will affect a firm's total revenue. Total revenue is equal to the price per unit multiplied by the number of units sold ($TR = P \times Q$). So as managers change the price of a product, sales generally vary. Good managers need to consider whether a price change will increase the firm's total revenue if their decisions are to improve firm performance. And whether a price change increases total revenue depends on the price elasticity of demand.

Suppose at the current price, demand for a product is price elastic; that is, the price elasticity of demand is less than -1 . In this situation, if the price is reduced, the percentage increase in quantity demanded is greater than the percentage reduction in price (this follows from the definition of elastic demand). That is, although all units are now being sold at a lower price, the increase in units sold because of

the lower price more than makes up for the slightly lower unit price. Hence total revenue increases. Similarly, if demand is elastic at the current price and a manager increases the price, total revenue will decrease. Although we are selling each unit at a higher price, the decrease in sales (because of the higher price) more than offsets the slightly higher price per unit.

If the demand for the product at a given price is price inelastic, the price elasticity of demand is greater than -1 . Following from our definition of inelastic demand, the percentage change in quantity is less than the percentage change in price. Hence if we increase price, total revenue will increase because the slightly higher price per unit sold will more than offset the decrease in units sold. If demand is inelastic and we decrease price, total revenue will decrease because the increase in units sold cannot offset the lower price per unit sold.

We can easily show this relationship formally. Let $TR = P \times Q$; then,

$$\begin{aligned}\frac{\Delta TR}{\Delta P} &= Q \frac{\Delta P}{\Delta P} + P \frac{\Delta Q}{\Delta P} \\ Q \frac{\Delta P}{\Delta P} + Q \frac{P}{Q} \frac{\Delta Q}{\Delta P} &= Q(1 + \eta) \\ \frac{\Delta TR / \Delta P}{Q} &= 1 + \eta\end{aligned}$$

QUANT OPTION

Although most students could hardly agree less, those who use calculus do find some measure of satisfaction in it. So here we derive the elasticity relationship using calculus

$$\begin{aligned}\frac{dTR}{dP} &= Q \frac{dP}{dP} + P \frac{dQ}{dP} \\ Q \frac{dP}{dP} + Q \frac{P}{Q} \frac{dQ}{dP} &= Q(1 + \eta) \\ \frac{dTR/dP}{Q} &= 1 + \eta\end{aligned}$$

If $\eta < -1$ (that is, elastic), $\frac{dTR/dP}{Q} < 0$ or $\frac{dTR}{dP} < 0$, so an increase in price will reduce TR . And if $\eta > -1$ (that is, inelastic), $\frac{dTR/dP}{Q} > 0$ or $\frac{dTR}{dP} > 0$, so an increase in price will increase TR .

PROBLEM SOLVED: Price Elasticity of Demand: Philip Morris

In 1993 Philip Morris cut cigarette prices by 18%. Its major competitor (RJ Reynolds) matched the price cut. Not surprisingly, the quantity sold of Philip Morris cigarettes increased (by 12.5%). In a June 13, 1994, article referring to the perils of a price cut, *Fortune* reported that Philip Morris profits fell by 25% as the result of a bad pricing strategy. Is there any evidence to determine whether this decision by Philip Morris managers decreased firm performance?

Although all the information is not available, we are not surprised by this result. We estimate the

price elasticity of demand for Philip Morris brands (including the iconic Marlboro Man) as revealed by the market:

$$\eta_p = \frac{\% \Delta Q}{\% \Delta P} = \frac{12.5\%}{-18\%} = -0.694$$

Demand is inelastic, so any drop in price should surely decrease firm revenue. Total revenue decreased and total costs increased (because more cigarettes were produced), so profit was destined to fall.

Then if $\eta < -1$ (that is, elastic), $\frac{\Delta TR / \Delta P}{Q} < 0$ or $\frac{\Delta TR}{\Delta P} < 0$, so an increase in price will reduce TR . And if $\eta > -1$ (that is, inelastic), $\frac{\Delta TR / \Delta P}{Q} > 0$ or $\frac{\Delta TR}{\Delta P} > 0$, so an increase in price will increase TR .

FUNDING PUBLIC TRANSIT

Consider an example in which we can predict market behavior based on estimates of elasticity measures. The fare (price) elasticity for public transportation in the United States is about -0.3 (that is, fairly inelastic). All transit systems in the United States lose money. Keeping the deficit under control is a constant battle because the typical subsidizers (federal, state, and local governments) are often reluctant to fund public transit (because of their own deficit problems). Can we identify which transit systems have the most difficult time getting public funding?

We can use our knowledge of price elasticity of demand. Managers of transit systems depend on two sources of revenue: ticket sales and public funding. They know that increasing fares will result in higher revenues. Most likely costs will also drop because less capital and labor are used. However, increasing fares reduces ridership and makes public transit less affordable to many. So managers of transit systems that do not receive enough public funding must increase their fares to balance their budgets (and they know their revenues will increase because of their inelastic demand).

DETERMINANTS OF THE OWN-PRICE ELASTICITY OF DEMAND

Table 2.4 shows the price elasticity of demand for selected products in the United States. Managers need to understand the factors that determine a product's price elasticity of demand:

1. The price elasticity of demand for a product depends heavily on the number and similarity of available substitute products. A product with many close

TABLE 2.4

Own Price Elasticities of Demand, Selected Goods, and Services from Global Locations

Good/Service	Elasticity	Good/Service	Elasticity
Agricultural products		Cigarettes (U.S.) ⁷	−1.107
Apples (U.S.) ¹	−1.159	Bread (U.K.) ³	−0.26
Potatoes (U.K.) ³	−0.13	Energy	
Oranges (U.S.) ²	−0.62	Gasoline—short run (Canada) ⁸	−0.01 to −0.2
Lettuce (U.S.) ²	−2.58	Gasoline—long run (Canada) ⁸	−0.4 to −0.8
Products from animals/fish		Transportation	
1 percent milk (U.S.) ⁵	−0.54 to −0.74	Domestic cars (U.S.) ⁹	−0.78
Cheese (U.K.) ³	−1.36	European cars (U.S.) ⁹	−1.09
Cheese (U.S.) ⁶	−0.595	Other manufactured goods	
Meat (China) ⁴	−0.06 to −0.18	Clothing and footwear (U.K./Ireland) ¹⁰	−0.94
Beef/veal (U.K.) ³	−1.45	Other goods (U.K./Ireland) ¹⁰	−0.85
Manufactured agricultural products		Services	
Beer and malt beverages (U.S.) ⁶	−2.83	Child care (North America) ¹¹	−0.570
Wine (U.K./Ireland) ⁷	−1.12	Government health care (Kenya) ¹²	−0.100
Wine and brandy (U.S.) ⁶	−0.198		

¹C. Elmore, Chapter 10, "Use of 2,4-D in Orchard, Vineyard, and Soft Fruit Production in the United States," *Phenoxy Herbicides*, December 20, 1998.

²D. Suits, "Agriculture," in Walter Adams and James Brock, eds., *The Structure of American Industry*, (10th ed.; Englewood Cliffs, NJ: Prentice-Hall, 2000).

³AEF116: 1.6: Major market response concepts and measures 1: The demand side and its elasticities.

⁴Millennium Institute, China Agricultural Project.

⁵A *Regional Economic Analysis of Dairy Compacts: Implications for Missouri Dairy Producers*, Section IV—"Economic Analysis of Dairy Compact, circa 1999."

⁶Emilo Pagoulatos and Robert Sorensen, "What Determines the Elasticity of Industry Demand," *International Journal of Industrial Organization*, Vol. 4, 1986.

⁷C. O'Donoghue, "Carbon Dioxide, Energy Taxes, and Household In-

come," Department of Statistics and Social Policy, London School of Economics, October 13, 1998.

⁸"Potential for Fuel Taxes to Reduce Greenhouse Gas Emissions in Transportation," Hagler Bailly Canada for Department of Public Works and Government Services, Hull, Quebec, June 11, 1990.

⁹P. McCarthy, "Market Price and Income Elasticities of New Vehicle Demands," *Review of Economics and Statistics*, Vol. 78(3), August 1996, 543–547.

¹⁰E. Brynjolfsson, "Some Estimates of the Contribution of Information Technology to Consumer Welfare," MIT Sloan School, Working Paper 3647-094, January 1994.

¹¹D. Chaplin et al., "The Price Elasticity of Child Care Demand: A Sensitivity Analysis."

¹²Section 4: "The Basics of Markets and Health Care Markets": Box 4.4: "Demand for Health Care in Kenya."

substitutes generally has elastic demand. If managers increase the product's price, consumers can easily switch to one of the several available substitutes. Conversely, if managers reduce the price of their product, they will see a significant increase in sales as consumers switch to their product. The extent to which a product has close substitutes largely depends on how well managers differentiate their product from similar ones.

2. The price elasticity of demand is also affected by a product's price relative to a consumer's total budget. Some claim that the demand for products like thimbles, rubber bands, and salt is quite inelastic, because the typical consumer spends only a very small fraction of her income on such goods. This is also why retail stores place items such as candy, soda, and magazines at the checkout counter. Because they are relatively inexpensive, consumers often buy them without thinking about the price. In contrast, products that command a larger percentage of the consumer's total budget tend to be more price elastic. Research has shown that when consumers consider purchasing items such as kitchen appliances or automobiles, they take the time to get several price quotes and gather information about brand attributes.

3. The price elasticity of demand for a product is also affected by the length of the period to which the demand curve pertains. For nondurable goods, demand is likely to be more elastic over a long period relative to a short period. This is because the longer the time period, the easier it is for consumers to substitute one good for another. If, for instance, the price of oil should decline relative to other fuels, the consumption of oil on the day after the price decline will probably increase very little. But over several years, people have an opportunity to react to the price decline in choosing types of home heating fuel; thus the price decline will have a greater effect on oil consumption than in the shorter period of one day. For durable goods, the opposite is true. Let's assume a consumer has just purchased a car. If the price of the car falls soon after the purchase, it is unlikely the consumer will run out and purchase another car; hence demand is inelastic.

THE STRATEGIC USE OF THE PRICE ELASTICITY OF DEMAND

Good managers not only display an avid interest in the price elasticity of demand for their products; they also take strategic actions to use the price elasticity to their benefit. Consider Table 2.5, which provides estimates of the price elasticity of demand for first-class, regular economy, and excursion air tickets between the United States and Europe. The price elasticity of demand for first-class air tickets is much lower in absolute value than for regular economy or excursion tickets, owing in part to the fact that the people who fly first class—often business travelers and relatively wealthy people—are unlikely to change their travel plans if moderate increases or decreases occur in the price of an air ticket. Airline executives study

STRATEGY SESSION: Elasticity in Use

Many of us are accustomed to see prices that end in 0s, 5s, or 9s. But prices that end in any integer are becoming more common. That's because firms are turning away from cost-plus pricing ("the pricing method still in use by the majority of manufacturers," IndustryWeek.com Leadership in Manufacturing) and turning toward strategic or value-based pricing. Profit margin improvements of as much as 21% are reported by Ralph Zuponicic, a managing partner of PricePoint Partners, one of many firms (such as SAP Khimetrics—an early entrant into this field) that now ply this market. Zuponicic notes that price changes are not usually huge (he states that typically a 1% price improvement can yield an 11% gain in profit margin for a firm with an 8% EBIT). Larry Robinson, director of Training and Development at PricePoint states that "adjusting prices for customer segments with varying sensitivity to value and price are the first steps to maximizing profits." This sensitivity is what we refer to in the text as price elasticity, and its precise role is shown in the text and later in this note.

Value-based pricing is one reason Wal-Mart has slashed prices (many to non-zero, five, or nine integers) and raised prices too. Longs Drug Stores (acquired by CVS) has also used such a model as has D'Agostino's (a New York-based grocery store chain). Not only do product prices go up or down from their

current levels (as specified by the model), but they also differ for the same product from location to location. Why? Because that above-mentioned price sensitivity differs in different market segments. PricePoint refers to an industrial rubber parts manufacturer where the prices before the application of value pricing ranged from \$8 to \$12 for the same product. After the application of value pricing, some prices increased, some decreased, and the manufacturer's profit margin on the item rose 21%.

In the retail sector, rather than marking up costs, benchmarking competitors' prices, or guessing, price optimization models use data-mining techniques. Scanned transactions from cash registers, responses to sales promotions, and the like are used to estimate an individual demand curve for each product in each store. Much of this modeling is based on airline yield management systems (see Chapter 9). The goal driving the modeling is to find the crossover point between driving sales and giving away margin unnecessarily. That is consultant-speak. Let's put it into economist-speak.

Airline yield management models attempt to equate marginal expected revenues for each fare class. For instance, suppose two fare classes, 1 and 2, exist. As shown on page 52, marginal revenue (MR) is equal to $MR = P[1 + (1/\eta)]$, so equating marginal revenues for classes 1 and 2 yields

these data carefully and price these classes of tickets differently. For example, because the price elasticity of demand for first-class air tickets is relatively low in absolute value, they price these tickets relatively high. In early 2012 a consumer could purchase an economy airline ticket for a round-trip between Philadelphia and Paris for under \$900. A first-class ticket would cost the consumer more than \$4,500.

Managers can also change the price elasticity of demand for their product. The most common way managers impact the price elasticity of their product is with differentiation strategies. Managers who successfully increase the differentiation

$$MR_1 = P_1[1 + (1/\eta_1)] = P_2[1 + (1/\eta_2)] = MR_2$$

Think about why a business with two different demand curves for the same product (such as business and leisure travelers for an airline seat, spring and summer demand for a bathing suit, location 1 and location 2 demand for Pampers) would want to equate the marginal revenues of these demand curves. If the marginal revenue from fare type 2 exceeds the marginal revenue from fare type 1, it would pay the airline to switch seats out of fare type 1 and into fare type 2. Because the costs of flying a person with fare type 1 is likely to be the same as the cost of flying a person with fare type 2, the airline can increase revenues while leaving costs the same by such a switch. Such a move must increase the airline's bottom line. (If the equated marginal revenues equal the marginal cost of moving a passenger, profits are not only improved, they are maximized.)

How would an elasticity model do this in retailing? Consider the optimal discounting of a product over time. The subscript 1 stands for the first time period, and the subscript 2 stands for the second time period. If the marginal revenue is higher in time period 2 than in period 1, you would want to shift some merchandise from period 1 to period 2 (or if 1 and 2 refer to stores, shift some product from store 1 to store 2).

Suppose that $\eta_1 = -2$ and $\eta_2 = -3$. Then

$$\begin{aligned} MR_1 &= P_1[1 + (1/-2)] = P_1[1 - (1/2)] = P_1/2 \\ &= P_2[1 + (1/-3)] = P_2[1 - (1/3)] \\ &= 2P_2/3 = MR_2 \end{aligned}$$

Or $P_2 = 0.75P_1$; that is, the optimal discount on the product would be to sell the good in time period 2 for 25% off the price from time period 1. Lowering price increases the quantity demanded (because demand curves have an inverse relationship between price and quantity), hence the term *driving sales*. But lowering the price too much or too little will not give the seller the optimal profit margins; that is, the profit margin that yields maximum profit. This can be done only where $MR_1 = MR_2 = (MC)$.

An article in the *Economist* ("The Price is Wrong") notes that supermarket chains "can quickly and easily track customers' elasticity"—how their buying habits change in response to a price rise or discount." Supermarkets such as D'Agostino's in New York and Dominicks in Chicago have used elasticity-based models to help them make pricing decisions.

Source: "Are You Getting Your Pricing Right or Leaving Money on the Table?" March 14, 2011, at www.industryweek.com/PrintArticle.aspx?ArticleID=24090.

of their product decrease (in absolute value) its price elasticity of demand. Simply put, differentiation strategies convince consumers the product is unique; hence it has fewer substitutes. Because consumers perceive fewer substitutes, they act as if they are more price inelastic. This gives managers more freedom to increase price because sales will fall less.

It is important for managers to understand that differentiation is not effective if consumers do not perceive it. Conversely, differentiation does not require tangible differences in products. For example, bleach is a commodity product; its chemical formula is well known. However, Clorox Bleach is able to command

TABLE 2.5

Elasticities of Demand for Air Tickets between the United States and Europe

Type of Ticket	Price Elasticity	Income Elasticity
First class	−0.45	1.50
Regular economy	−1.30	1.38
Excursion	−1.83	2.37

Source: J. Cigliano, "Price and Income Elasticities for Airline Travel: The North Atlantic Market," *Business Economics*, September 1980.

a retail price that is 300% higher than other brands of bleach. Most remarkably, this brand has been able to command such a high premium over several decades.

TOTAL REVENUE, MARGINAL REVENUE, AND PRICE ELASTICITY

We want to look more closely at the effect of the price elasticity of demand on a firm's total revenue. To a good's producer, the total amount of money paid by consumers equals the firm's revenue. Therefore, to the Toyota Motor Company, the total amount consumers spend on its cars is its total revenue. Suppose the demand curve for a firm's product is linear; that is,

$$P = a - bQ \quad (2.11)$$

where a is the intercept on the price axis and b is the slope (in absolute terms), as shown in panel A of Figure 2.7. Thus the firm's total revenue equals

$$TR = PQ = (a - bQ)Q = aQ - bQ^2 \quad (2.12)$$

Marginal revenue The incremental revenue earned from selling the n th unit of output.

An important concept to managers is that of **marginal revenue**, which is the incremental revenue earned from selling the n th unit of output. As we will see, managers must understand marginal revenue to maximize the firm's profit. Because this concept is central to firm performance, we need to understand how the price elasticity of demand affects it. In the present case,

$$\begin{aligned}
 MR &= \frac{\Delta TR}{\Delta Q} \\
 &= \frac{\Delta(aQ - bQ^2)}{\Delta Q} \\
 &= a - 2bQ
 \end{aligned} \quad (2.13)$$

STRATEGY SESSION: Elasticity in Specialty Drugs

Specialty drugs are designed to treat illnesses that have major impacts on patients' lives, are difficult to treat, and don't affect a large number of people; these are very expensive to produce.

One such disease is multiple sclerosis (MS). It affects slightly over two million people worldwide, and a year's worth of drug treatment can approach \$50,000.

A new oral medication, Novartis AG's Gilenya, has recently received FDA approval. Its yearly price tag is \$48,000. MS sufferers are looking forward to the drug because its substitutes require injections, are much harder to self-administer, and can have nasty side effects. The side effects of Gilenya are not completely known, but its ease of use is creating optimism.

The makers of injectibles are sensing that the demand for their product will decline. A recent survey of neurologists by a Citigroup analyst showed that all other MS drugs will suffer from declines in sales in the next year. Some expect Gilenya will be the market leader by 2017.

So what are the producers of injectibles doing? They are raising their prices. They rose 39% last year (2010 to 2011). Analysts say that charging more for the old treatment is a way to "keep revenues steady as sales erode." Teva and Biogen Idec are two companies seeking to generate more revenue before

Gilenya wins significant market share, according to the source article.

Under what conditions can the injectibles achieve their goals? For revenues to be constant as prices rise and quantity falls, the elasticity must be unity, that is $\eta = -1$, and for revenues to increase when prices increase and quantity falls, the demand must be inelastic. Given the nature of the disease, it's not too difficult to believe that demand could be inelastic. Once one has found an injectible that "works," it's likely that the user would stick with it.

And maybe hope is on the way. There are three other oral MS drugs in the testing stages that may be a couple of years away. Hopefully, competition among the producers can bring the price down. However, if the drugs differ with respect to their effectiveness and side effects, the product differentiation could still lead to significant pricing power for each version. In addition, even with a homogenous product, competition driving the price down to cost could still lead to a hefty price for the drug given the high costs of production.

Source: Eva von Schaper and Naomi Kresge, "Novartis's \$48,000 Pill Spurs US Price Increase for MS Drugs," *Bloomberg News*, March 22, 2011, at www.bloomberg.com/news/2011-03-21/novartis-s-48-000-pill-spurs-u-s-price-increases-for-ms-drugs.html.

which is also shown in panel A of Figure 2.7. Comparing the marginal revenue curve with the demand curve, we see that while both have the same intercept on the vertical axis (this intercept being a) the slope of the marginal revenue curve is twice that of the demand curve.

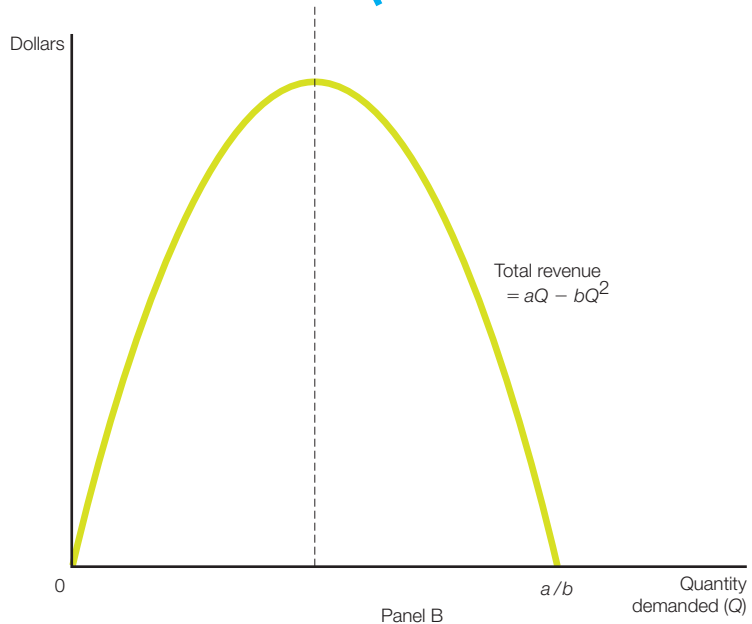
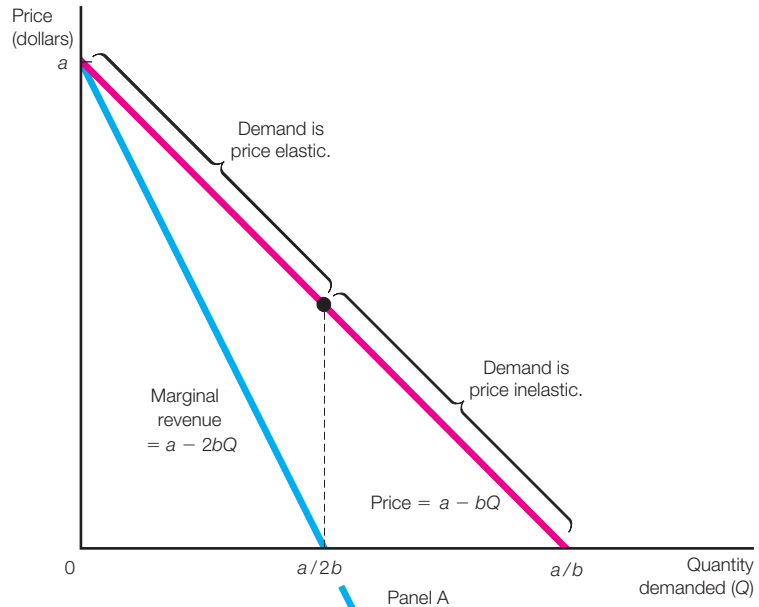
According to the definition in equation (2.7), the price elasticity of demand, η , equals $(\Delta Q / \Delta P)(P / Q)$. Because $\Delta Q / \Delta P = -1/b$ and $P = a - bQ$, it follows, in this case, that

$$\eta = \left(\frac{-1}{b} \right) \frac{a - bQ}{Q} \quad (2.14)$$

FIGURE 2.7

Relationship between Price Elasticity, Marginal Revenue, and Total Revenue

If demand is price elastic, marginal revenue is positive and increases in quantity result in higher total revenue. If demand is price inelastic, marginal revenue is negative and increases in quantity result in lower total revenue.



QUANT OPTION

We can see this relationship more clearly using calculus:

$$\begin{aligned} MR &= \frac{dTR}{dQ} \\ &= \frac{d[aQ - bQ^2]}{dQ} \\ &= a - 2bQ \end{aligned}$$

Therefore, whether η is greater than, equal to, or less than -1 depends on whether Q is greater than, equal to, or less than $a/2b$. As shown in Figure 2.7, demand is price elastic if $Q < a/2b$; it is of unitary elasticity if $Q = a/2b$; and it is price inelastic if $Q > a/2b$.

Panel B in Figure 2.7 plots the firm's total revenue against the quantity demanded of its product. Remember that marginal revenue is the incremental revenue earned from selling the next unit of output. As long as marginal revenue is positive, an increase in sales raises total revenue. However, at outputs where the incremental revenue is negative, total revenue will decrease. Some may ask, how can incremental revenue be negative? If a firm sells one more unit, it must receive a positive revenue from the person it sells to (unless it pays the person to take the good). But think about the effect on total revenue of that last unit sold.

STRATEGY SESSION: Beer—Elastic or Inelastic?

If you drink beer, you'd better hope your pants have some elasticity. But how about demand elasticity? InBev (brewers of Budweiser, Stella Artois, and Beck's) recently announced that due to high unemployment of young males in the United States, quantity demanded of their products fell by 0.4%, while prices rose. The direction of the change is no surprise given our knowledge of demand curves. But they also reported that total revenue rose by a whopping 5.6%. What does this tell us? That the quantity demanded stayed very high and in percentage terms

dropped less than the percentage increase in price. Or in economic terms: The demand for their product is inelastic. But how inelastic? From the text, we have that $\Delta TR/\Delta Q = P[1 + (1/\eta)]$. Multiplying both sides by Q/TR (or $Q/TR = Q/PQ = 1/P$) gives $(\Delta TR/TR)/(\Delta Q/Q) = \% \Delta TR/\% \Delta Q = 5.6/0.4 = -14 = [1 + (1/\eta)]$ or $1/\eta = -15$ or $\eta = 1/-15 = -0.067$. This is very inelastic.

Source: "InBev Price Hike Outweighs Lower Volumes," at www.cnbc.com/id/42891391/.

If managers need to reduce the price to sell that last unit, they reduce the price of all the units sold. Hence selling one more unit can cause total revenue to be lower because the manager has reduced the price on all the units sold. Also, producing one more unit increases total costs. So if total revenue decreases and total cost increases, the manager is moving further from a profit-maximizing strategy. The important thing to remember is that managers do not want to produce at an output level where marginal revenue is negative.

Another thing to note about Figure 2.7 is that, at quantities where demand is price elastic, marginal revenue is positive; at quantities where it is of unitary elasticity, marginal revenue is zero; and at quantities where it is price inelastic, marginal revenue is negative. This is no accident. In general, whether or not the demand curve is linear, this is the case. To see why, recall that by definition,

$$MR = \frac{\Delta TR}{\Delta Q}$$

Because total revenue equals price times quantity, it follows that

$$MR = \frac{\Delta(PQ)}{\Delta Q}$$

We can transform this into

$$MR = P \frac{\Delta Q}{\Delta Q} + Q \frac{\Delta P}{\Delta Q}$$

Because $\Delta Q / \Delta Q = 1$,

$$\begin{aligned} MR &= P + Q \frac{\Delta P}{\Delta Q} \\ &= P \left[1 + \left(\frac{Q}{P} \right) \left(\frac{\Delta P}{\Delta Q} \right) \right] \end{aligned}$$

And because the definition of the price elasticity of demand implies that $(Q/P)(\Delta P / \Delta Q) = 1/\eta$,

$$MR = P \left(1 + \frac{1}{\eta} \right) \tag{2.15}$$

Equation (2.15) shows that if $\eta < -1$, marginal revenue is positive; if $\eta > -1$, marginal revenue is negative; and if $\eta = -1$, marginal revenue is zero. In later chapters we use equation (2.15) repeatedly. Managers will find lots of uses for it. One way they use it is to estimate the value of marginal revenue. For example, what is the marginal revenue if the product price is \$10 and the price elasticity of demand is -2 ? On the basis of equation (2.15), it equals $10(1 - 1/2) = \$5$.

QUANT OPTION

Equation (2.15) is so famous that we think it deserves a formal derivation

$$\begin{aligned}
 MR &= \frac{dTR}{dQ} \\
 MR &= \frac{d(PQ)}{dQ} \\
 MR &= P \frac{dQ}{dQ} + Q \frac{dP}{dQ} \\
 MR &= P + Q \frac{dP}{dQ} \\
 &= P \left[1 + \left(\frac{Q}{P} \right) \left(\frac{dP}{dQ} \right) \right] \\
 MR &= P \left(1 + \frac{1}{\eta} \right)
 \end{aligned}$$

THE INCOME ELASTICITY OF DEMAND

Price is not the only factor managers need to consider in predicting how many units they will sell. An important factor over which managers have little control is the level of consumer income. A consumer with relatively more money to spend is more likely to purchase various goods. For example, more cognac is sold on a per capita basis in cities populated by consumers with higher income levels, like New York City, than in cities like Fargo, North Dakota, where consumers have lower income levels. Managers need to understand the sensitivity of demand to changes in consumer incomes.

The **income elasticity of demand** for a particular good is defined as the percentage change in quantity demanded resulting from a 1% change in consumers' income. More precisely, it equals

$$\eta_I = \left(\frac{\Delta Q}{\Delta I} \right) \left(\frac{I}{Q} \right) \quad (2.16)$$

where Q is quantity demanded and I is consumers' income. For most products, the income elasticity of demand is positive. That is, when consumer incomes increase, they buy more of the product. Such goods are called *normal goods*. However, some products, called *inferior goods*, have negative income elasticities. For these goods, quantity demanded moves opposite to consumer incomes. When incomes increase, quantity demanded decreases; when incomes decrease, quantity demanded increases. Two examples of inferior goods in the United States are

Income elasticity of demand The percentage change in quantity demanded resulting from a 1% change in consumers' income.

hamburgers and public transportation. When consumers' income increases, they generally consume fewer hamburgers. It is not that they eat less beef, but they eat steaks instead of burgers.

Managers must understand the impact of the income elasticity of demand on sales. Variance in sales is common in products with high income elasticities as sales react to the various stages of the business cycle. When the economy is expanding, products with high income elasticities will enjoy a significant increase in sales. However, when the economy enters a recession, these same products will see a significant decrease in sales. Though managers selling products with high income elasticities can do little to influence the business cycle, they can plan ahead to lessen the negative impact of economic fluctuations. For example, they can try to shift their cost structure toward fewer fixed costs and greater variable costs. One way of accomplishing this is by leasing capital goods instead of buying them. And in times of economic expansion, they must prepare for the probable significant increase in sales.

In forecasting the long-term growth of the quantity demanded for many major products, the income elasticity of demand is of key importance. According

STRATEGY SESSION: Estimating the Demand for Amtrak Rail Passenger Business

Amtrak (the national passenger railroad in the United States) has used an aggregate demand model to forecast system-wide passenger revenues. The first part of its model was a multivariate linear regression that forecast its system-wide passenger miles—the dependent variable (a passenger mile is one passenger moved one mile). Explanatory variables were disposable personal income; Amtrak's average fare; the ratio of Amtrak's fare to the airlines' average fare; retail gasoline prices; dummy variables to reflect such events as weather, holidays, strikes, and derailments; and dummy variables to reflect seasonal variation.

The most important determinant of rail passenger miles was disposable personal income (a proxy variable for the strength of the U.S. economy). From the regression, a 1% increase in disposable personal income was expected to yield a 1.8% increase in system-wide passenger miles.

1. Explain the rationale for each explanatory variable appearing in the model and what sign (positive or negative) you expect for its regression coefficient.
2. How did Amtrak use this forecast of system-wide passenger miles to obtain its estimate of system-wide passenger revenues?
3. What is Amtrak's estimate of income elasticity of demand for passenger service (based on the income level of the economy as a whole)? If U.S. disposable personal income per capita increases from \$27,000 to \$28,000, what's your prediction for the increase in train passenger miles in the new situation (all other independent variables remaining constant)? Will this prediction be 100% accurate? Why or why not?
4. Even though disposable personal income rose from 2000 through 2008 (on a real income basis by 13.26%), Amtrak's market share of U.S. passenger miles traveled has remained virtually constant (0.1105% in 2000 and 0.1120% in 2008). Can you explain this?

to studies done by the U.S. Department of Agriculture, the income elasticity of demand for milk is about 0.5, which means a 1% increase in disposable income is associated with about a 0.5% increase in the quantity demanded of milk. But in a study done in Britain, the income elasticity of bread was about -0.17 , which means that a 1% increase in disposable income is associated with about a -0.17% decrease in the quantity demanded of bread. Table 2.5 shows that the income elasticity of demand for first-class air tickets between the United States and Europe is 1.5, which means that a 1% increase in disposable income is associated with a 1.5% increase in the quantity demanded of such tickets. Table 2.6 shows the income elasticity of demand for other commodities across the world. In measuring income elasticities, income can be defined as the aggregate income of consumers (as in Table 2.6) or as per capita income (as in the next section), depending on the circumstances.

CROSS-PRICE ELASTICITIES OF DEMAND

In addition to price and income, another factor influencing the quantity demanded of a product is the prices of rivals. Holding constant the product's own price (as well as the level of income) and allowing the price of another product to vary may result in important effects on the quantity demanded of the product in question.

TABLE 2.6

Income Elasticity of Demand, Selected Commodities, Global

Good	Elasticity	Good	Elasticity
Agricultural products		Cream (U.S.) ³	+1.72
Grain (China) ¹	-0.12 to $+0.15$	Eggs (U.K.) ²	-0.21
Potatoes (U.K.) ²	-0.32	Eggs (U.S.) ³	$+0.57$
Potatoes (U.S.) ³	$+0.15$	Processed food products	
Oranges (U.S.) ³	$+0.83$	Bread (U.K.) ²	-0.17
Apples (U.S.) ³	$+1.32$	Other cereal products (U.K.) ²	$+0.18$
Lettuce (U.S.) ³	$+0.88$	Automobiles	
Animal products		Domestic cars (U.S.) ⁴	$+1.62$
Meat (China) ¹	$+0.1$ to $+1.2$	European cars (U.S.) ⁴	$+1.93$
Milk (U.K.) ²	$+0.05$	Asian cars (U.S.) ⁴	$+1.65$
Milk (U.S.) ³	$+0.50$		

¹Millennium Institute, China Agricultural Project.

²AEF116: 1.6 "Major Market Response Concepts and Measures. 1: The Demand Side and Its Elasticities."

³D. Suits, "Agriculture," in *The Structure of American Industry*, ed. Adams and Brock.

⁴P. McCarthy, "Market Price and Income Elasticities of New Vehicle Demands."

Cross-price elasticity of demand

The percentage change in the quantity demanded of one good resulting from a 1% change in the price of another good.

By observing these effects, we can classify pairs of products as substitutes or complements, and we can measure how close consumers perceive the relationship (either substitute or complementary). Consider products X and Y . If Y 's price goes up, what is the effect on Q_X , the quantity of X demanded? The **cross-price elasticity of demand** is defined as the percentage change in the quantity demanded of good X resulting from a 1% change in the price of good Y

$$\eta_{XY} = \left(\frac{\Delta Q_X}{\Delta P_Y} \right) \left(\frac{P_Y}{Q_X} \right) \quad (2.17)$$

Goods X and Y are classified as substitutes if the cross-price elasticity of demand is positive. For instance, an increase in the price of wheat, when the price of corn remains constant, tends to increase the quantity of corn demanded; therefore, η_{XY} is positive, and wheat and corn are classified as substitutes. On the other hand, if the cross-price elasticity of demand is negative, goods X and Y are classified as complements. For example, an increase in the price of software tends to decrease the purchase of laptop computers when the price of laptops remains constant; therefore, η_{XY} is negative, and software and laptops are classified as complements. If the cross-price elasticity of two products is around zero, then the products have independent demand levels. For example, if the price of butter increases, the demand for airline tickets remains constant.

PROBLEM SOLVED: Income Elasticity of Demand

In a previous section we learned how to calculate the price elasticity of demand based on a product's demand function. Here we see how to calculate the income elasticity of demand. Suppose the demand function for a product is

$$Q_X = 1,000 - 0.2P_X + 0.5P_Y + 0.04I$$

where Q_X is the quantity demanded of good X , P_X is the price of good X , P_Y is the price of good Y , and I is per capita disposable income. The income elasticity of demand is

$$\begin{aligned} \eta_I &= \left(\frac{\Delta Q}{\Delta I} \right) \left(\frac{I}{Q} \right) \\ &= 0.04 \frac{I}{Q} \end{aligned}$$

If $I = 10,000$ and $Q = 1,600$

$$\eta_I = 0.04 \left(\frac{10,000}{1,600} \right) = 0.25$$

The income elasticity of demand equals 0.25, which means that a 1% increase in per capita disposable income is associated with a 0.25% increase in the quantity demanded of product X .

To illustrate the calculation of cross-price elasticities, suppose once again the demand function for our product is

$$Q_X = 1,000 - 0.2P_X + 0.5P_Y + 0.04I$$

where Q_X is the quantity demanded of product X , P_X is the price of X , P_Y is the price of product Y , and I is per capita disposable income. The cross-price elasticity of demand between products X and Y is

$$\begin{aligned}\eta_{XY} &= \left(\frac{\Delta Q_X}{\Delta P_Y} \right) \left(\frac{P_Y}{Q_X} \right) \\ &= 0.5 \frac{P_Y}{Q_X}\end{aligned}$$

Although the value of the cross-price elasticity depends on the values of P_Y and Q_X , the goods are always substitutes because η_{XY} is positive, regardless of the values of P_Y and Q_X . If $P_Y = 500$ and $Q_X = 2,000$,

$$\eta_{XY} = 0.5 \left(\frac{500}{2,000} \right) = 0.125$$

The cross-price elasticity of demand is of fundamental importance to managers because they continually must anticipate what will happen to their own sales if rivals change their prices. To do so, they need information concerning the cross-price elasticities of demand. Table 2.7 shows the cross-price elasticities of demand for selected pairs of commodities.

The measure is also frequently used by antitrust authorities to evaluate proposed mergers. A high cross-price elasticity measure between products X and Y can cause concern that a merger between the producers of X and Y might result in consumers experiencing higher prices and fewer brand choices. A highly negative cross-price elasticity measure ($-\eta_{xy}$) signifies that the products are strong complements. Here the authorities may be concerned a merger between the products might lead to excessive control of the supply chain. That is, the merged firm may refuse to sell the intermediate product to other producers.

THE ADVERTISING ELASTICITY OF DEMAND

Although the price elasticity, income elasticity, and cross-price elasticities of demand are the most frequently used elasticity measures, they are not the only ones. For example, managers sometimes find it useful to calculate the **advertising elasticity of demand**. Suppose the demand function for a particular firm's product is

$$Q = 500 - 0.5P + 0.01I + 0.82A$$

where Q is the quantity demanded of the product, P is its price, I is per capita disposable income, and A is the firm's advertising expenditure. The advertising

Advertising elasticity of demand
The percentage change in the quantity demanded of the product resulting from a 1% change in the advertising expenditure.

TABLE 2.7

Cross-Price Elasticity of Demand, Selected Pairs of Commodities, Global

Change of Price of Good	Change of Quantity of Good	Cross-Price Elasticity
European/Asian cars	U.S. domestic cars	+0.28 ¹
European/U.S. domestic cars	Asian cars	+0.61 ¹
U.S. domestic/Asian cars	European cars	+0.76 ¹
Australian public transit	Australian auto ownership	+0.1 to +0.3 ²
Irish coal	Irish natural gas	+0.4 ³
Irish coal	Irish oil	+0.7 ³
Kenyan government-provided health care	Mission- or private sector-provided health care in Kenya	+0.023 ⁴
U.S. durum wheat	U.S. hard red spring wheat	+0.04 ⁵
U.S. hard red winter wheat	U.S. white wheat	+1.80 ⁵
U.K. beef/veal	U.K. pork	0.00 ⁶
U.K. mutton/lamb	U.K. beef/veal	+0.25 ⁶

¹P. McCarthy, "Market Price and Income Elasticities of New Vehicle Demand."

²J. Luk and S. Hepburn, "A Review of Australian Travel Demand Elasticities," Working Document No. TE 93/004, 1993, Australian Road Research Board.

³Competition Authority Decision of 30 January 1998, relating to a proceeding under Section 4 of the Competition Act 1991: Notification No. CA/15/97—Statoil Ireland Ltd./Clare Oil Company Ltd.—Share Purchase Agreement

and Service Employment Agreement. Decision No. 490.

⁴"Section 4: The Basics of Markets and Health Care Markets: Box 4.4: Demand for Health Care in Kenya."

⁵Wheat Yearbook, March 30, 1998, Economic Research Services, U.S. Department of Agriculture, Washington, DC 20036-5831.

⁶AEF116: 1.6: "Major Market Response Concepts and Measures. 1: The Demand Side and Its Elasticities."

elasticity is defined as the percentage change in the quantity demanded of the product resulting from a 1% change in advertising expenditure. More precisely, it equals

$$\eta_A = \left(\frac{\Delta Q}{\Delta A} \right) \left(\frac{A}{Q} \right) \quad (2.18)$$

In this case, because $\Delta Q / \Delta A = 0.82$,

$$\eta_A = 0.82 \frac{A}{Q}$$

If A/Q , the amount of advertising per unit of the product demanded, is \$2,

$$\eta_A = 0.82(2) = 1.64$$

This useful elasticity tells managers that a 1% increase in advertising expenditure results in a 1.64% increase in the quantity demanded. In later chapters we will see how information of this sort is used to help guide managerial decisions.

THE CONSTANT-ELASTICITY AND UNITARY ELASTIC DEMAND FUNCTION

In this chapter, we generally assume the demand function is linear. That is, the quantity demanded of a product is assumed to be a linear function of its price, the prices of other goods, consumer income, and other variables. Another mathematical form frequently used is the **constant-elasticity demand function**. If the quantity demanded (Q) depends only on the product's price (P) and consumer income (I), this mathematical form is

$$Q = aP^{-b_1/b_2} \quad (2.19)$$

Therefore, if $a = 200$, $b_1 = 0.3$, and $b_2 = 2$,

$$Q = 200P^{-0.3/2}$$

An important property of this type of demand is that the price elasticity of demand equals $-b_1$, regardless of the value of P or I . (This accounts for it being called the constant-elasticity demand function.)

The constant-elasticity demand function is often used by managers and economists for several reasons. First, in contrast to the linear demand function, this mathematical form explicitly recognizes the effect of price on quantity demanded depends on income level and that the effect of income on quantity demanded depends on price. The multiplicative relationship in equation (2.19) is often more realistic than the additive relationship in equation (2.1). Second, like the linear demand function, the constant-elasticity demand function is relatively easy to estimate.

If the demand is of unitary elasticity (the price elasticity of demand equals -1), an increase or decrease in price has no effect on the amount spent on the good. As an illustration, consider the case shown in Figure 2.8. The demand curve shown is a rectangular hyperbola, which means that

$$Q = \frac{m}{P} \quad (2.20)$$

Constant-elasticity demand function Mathematical form that always yields the same elasticity, regardless of the product's price and the consumers' income.

QUANT OPTION

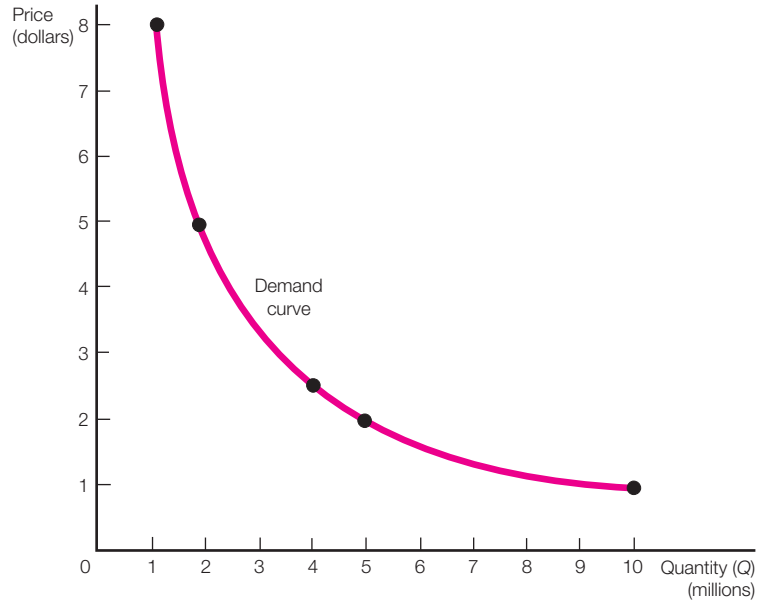
Here is the formal derivation of the constant elasticity.

$$\eta = (P/Q)[\partial Q/\partial P] = (P/aP^{-b_1/b_2})[-ab_1P^{b_1-1/b_2}] = P[-b_1P^{-1}] = -b_1$$

FIGURE 2.8

Demand Curve with Unitary Elasticity at All Points

The demand curve is a rectangular hyperbola if the price elasticity of demand is always -1 .



where Q is product demand, P is the price, and m is a constant. This type of demand is of unitary elasticity at all points. Hence changes in price have no effect on the total amount spent on the product. It is evident from equation (2.20) that, regardless of the price, the total expenditure for the product will be m (\$10 million in Figure 2.8).

SUMMARY

1. The market demand curve for a product shows how much of the product is demanded at each price. The market demand curve shifts in response to changes in tastes, incomes, the prices of other products, advertising, and the size of the population.
2. The market demand function for a product is an equation showing how the quantity demanded depends on the product's price, the incomes of consumers, the prices of other products, advertising expenditure, and additional factors. Holding all factors other than the product's price constant, we can draw the market demand curve for the product from the market demand function. Mar-

ket demand functions are formulated for individual firms as well as for entire industries.

3. The own-price elasticity of demand is the percentage change in quantity demanded resulting from a 1% change in price; more precisely, it equals $(\Delta Q / \Delta P)(P / Q)$. Whether a price increase (or decrease) results in an increase in the total amount spent by consumers on a product depends on the own-price elasticity of demand.

4. Marginal revenue is the change in total revenue resulting from a one-unit increase in quantity. Marginal revenue equals $P(1 + 1/\eta)$, where P is price and η is the own-price elasticity of demand.

5. The own-price elasticity of demand for a product tends to be elastic if the product has many close substitutes. Also, it often tends to be more elastic in the long run than in the short run. It is sometimes asserted a product's demand is relatively price inelastic if the product accounts for a very small percentage of the typical consumer's budget, but this need not be the case.

6. The income elasticity of demand is the percentage change in quantity demanded resulting from a 1% change in consumer income; that is, it equals $(\Delta Q / \Delta I)(I / Q)$, where I is the income of consumers. The income elasticity of demand may be positive or negative. Like the price elasticity of demand, it is of major importance in forecasting the long-term growth in the quantity demanded for products.

7. The cross-price elasticity of demand is the percentage change in the quantity demanded of product X resulting from a 1% change in the price of product Y ; in other words, it equals $(\Delta Q_X / \Delta P_Y)(P_Y / Q_X)$. If X and Y are substitutes, it is positive; if they are complements, it is negative. This elasticity is important for managers because they must understand and forecast the effects of changes in other firms' prices on their own firm's sales.

8. If a demand curve is linear, the own-price elasticity of demand varies from point to point on the demand curve. As price approaches zero, the own-price elasticity of demand also approaches zero. As quantity demanded approaches zero, the own-price elasticity approaches negative infinity. In contrast, for a constant-elasticity demand function, the own-price elasticity of demand is the same regardless of the product's price. Both linear demand functions and constant-elasticity demand functions are used frequently by managers and managerial economists.

PROBLEMS

1. The Dolan Corporation, a maker of small engines, determines that in 2012 the demand curve for its product is

$$P = 2,000 - 50Q$$

where P is the price (in dollars) of an engine and Q is the number of engines sold per month.

- a. To sell 20 engines per month, what price would Dolan have to charge?
 - b. If managers set a price of \$500, how many engines will Dolan sell per month?
 - c. What is the price elasticity of demand if price equals \$500?
 - d. At what price, if any, will the demand for Dolan's engines be of unitary elasticity?
2. The Johnson Robot Company's marketing managers estimate that the demand curve for the company's robots in 2012 is

$$P = 3,000 - 40Q$$

where P is the price of a robot and Q is the number sold per month.

- a. Derive the marginal revenue curve for the firm.
 - b. At what prices is the demand for the firm's product price elastic?
 - c. If the firm wants to maximize its dollar sales volume, what price should it charge?
3. After a careful statistical analysis, the Chidester Company concludes the demand function for its product is

$$Q = 500 - 3P + 2P_r + 0.1I$$

where Q is the quantity demanded of its product, P is the price of its product, P_r is the price of its rival's product, and I is per capita disposable income (in dollars). At present, $P = \$10$, $P_r = \$20$ and $I = \$6,000$.

- a. What is the price elasticity of demand for the firm's product?
 - b. What is the income elasticity of demand for the firm's product?
 - c. What is the cross-price elasticity of demand between its product and its rival's product?
 - d. What is the implicit assumption regarding the population in the market?
4. The Haas Corporation's executive vice president circulates a memo to the firm's top management in which he argues for a reduction in the price of the firm's product. He says such a price cut will increase the firm's sales and profits.
 - a. The firm's marketing manager responds with a memo pointing out that the price elasticity of demand for the firm's product is about -0.5 . Why is this fact relevant?
 - b. The firm's president concurs with the opinion of the executive vice president. Is she correct?
 5. Managers of the Hanover Manufacturing Company believe the demand curve for its product is

$$P = 5 - Q$$

where P is the price of its product (in dollars) and Q is the number of millions of units of its product sold per day. It is currently charging \$1 per unit for its product.

- a. Evaluate the wisdom of the firm's pricing policy.
- b. A marketing specialist says that the price elasticity of demand for the firm's product is -1.0 . Is this correct?
6. On the basis of historical data, Richard Tennant has concluded, "The consumption of cigarettes is . . . [relatively] insensitive to changes in price. . . . In contrast, the demand for individual brands is highly elastic in its response to price. . . . In 1918, for example, Lucky Strike was sold for a short time at a higher retail price than Camel or Chesterfield and rapidly lost half its business."
 - a. Explain why the demand for a particular brand is more elastic than the demand for all cigarettes. If Lucky Strike raised its price by 1% in 1918, was the price elasticity of demand for its product greater than -2 ?
 - b. Do you think that the demand curve for cigarettes is the same now as it was in 1918? If not, describe in detail the factors that have shifted the demand curve and whether each has shifted it to the left or right.
7. According to S. Sackrin of the U.S. Department of Agriculture, the price elasticity of demand for cigarettes is between -0.3 and -0.4 , and the income elasticity of demand is about 0.5 .
 - a. Suppose the federal government, influenced by findings that link cigarettes and cancer, were to impose a tax on cigarettes that increased their price by 15%. What effect would this have on cigarette consumption?
 - b. Suppose a brokerage house advised you to buy cigarette stocks because if incomes were to rise by 50% in the next decade, cigarette sales would be likely to spurt enormously. What would be your reaction to this advice?
8. A survey of major U.S. firms estimates on average, the advertising elasticity of demand is only about 0.003 . Doesn't this indicate that managers spend too much on advertising?
9. The McCauley Company hires a marketing consultant to estimate the demand function for its product. The consultant concludes that this demand function is

$$Q = 100P^{-3.1}I^{2.3}A^{0.1}$$

where Q is the quantity demanded per capita per month, P is the product's price (in dollars), I is per capita disposable income (in dollars), and A is the firm's advertising expenditures (in thousands of dollars).

- a. What is the price elasticity of demand?
- b. Will price increases result in increases or decreases in the amount spent on McCauley's product?
- c. What is the income elasticity of demand?
- d. What is the advertising elasticity of demand?
- e. If the population in the market increases by 10%, what is the effect on the quantity demanded if P , I , and A are held constant?

10. The Schmidt Corporation estimates that its demand function is

$$Q = 400 - 3P + 4I + 0.6A$$

where Q is the quantity demanded per month, P is the product's price (in dollars), I is per capita disposable income (in thousands of dollars), and A is the firm's advertising expenditures (in thousands of dollars per month). Population is assumed to be constant.

- a. During the next decade, per capita disposable income is expected to increase by \$5,000. What effect will this have on the firm's sales?
- b. If Schmidt wants to raise its price enough to offset the effect of the increase in per capita disposable income, by how much must it raise its price?
- c. If Schmidt raises its price by this amount, will it increase or decrease the price elasticity of demand? Explain. Make sure your answers reflect the fact that elasticity is a negative number.

CONSUMER BEHAVIOR AND RATIONAL CHOICE

We discussed market demand in Chapter 2. But as we will soon see, market demand for a product is the aggregate of individual demand for that product. So managers need to understand how individuals choose products. As you read this chapter, think about how you make decisions; certainly we all purchase much in our lives. More important, you need to understand how the variables controlled and directed by managers (prices, advertising, etc.) influence consumer choice.

Our model for consumer behavior is part of a larger whole. All of us make decisions every day. Most are decided with little effort: Either the choice is obvious or the impact is limited, so deep thought is not warranted. Occasionally we encounter decisions that require more thought and have high impact; in these situations we are likely to think harder about possible choices and their consequences. But whether we make a snap judgment or a systematic analysis, we are governed by an internal classification scheme that tells us we prefer one choice over another. Without preference ordering, we are reduced to random choices.

This chapter shows how economists model consumer purchase decisions. In later chapters we examine how individuals make decisions under risk (Chapter 14) and when they possess asymmetric information (Chapters 15 and 16).

Although some students may fret about the usefulness of economic principles in the business world, they cannot deny the usefulness of our decision models. It should be obvious that managers constantly face limited budgets to allocate across different uses. You as a consumer face this decision daily. For a problem of this sort,

LEARNING OBJECTIVES

Indifference Curves

The Marginal Rate of Substitution

The Concept of Utility

The Budget Line

The Equilibrium Market Bundle

Maximizing Utility: A Closer Look

Corner Solutions

How Managers Can Strategically Influence Consumer Choices

Deriving the Individual Demand Curve

Deriving the Market Demand Curve

Consumer Surplus

Summary

Problems

the economist's model of consumer behavior provides some guidelines. And good managers understand that they can take actions to influence consumer choice. This is the idea underlying the use of marketing, pricing, and distributional strategies.

In subsequent chapters we explain how this decision framework is applied to help improve managerial decision making. For now, you may want to try Problem 11 at the end of this chapter. See if you can figure out how this model sheds light on a state's choice between mass transit and highways. (The answer is provided at the end of the book.)

In examining how consumers choose, we initially assume a consumer is rational and wishes to maximize his or her well-being. That is, to the best of their knowledge, consumers do not make choices that cause them harm. A consumer's well-being is a function of the goods she or he chooses to purchase. However, this well-being is not unconstrained. If it were, we would see many more people driving expensive cars like a Porsche or a Bentley. Purchases are constrained by the income level of the consumer. A rational consumer maximizes his or her well-being given the prices of goods, personal tastes and preferences for goods, and income. We formally model this behavior by developing the concepts of utility functions, indifference curves, and budget lines. Using them, we derive the consumer's demand curve for products and show how demand shifts when income changes.

INDIFFERENCE CURVES

To clarify important ideas, we initially assume consumers can purchase only food products and clothing products. All the implications we discuss apply to the more complex setting of the world. Consumer choice is modeled as a series of indifference curves. An **indifference curve** contains points representing market bundles among which the consumer is indifferent. To illustrate, consider Jennifer Popovich, a consumer in South Pasadena, California. Certain market bundles—that is, combinations of food and clothing—are equally desirable to her. For example, she may have a hard time choosing between a market bundle containing 50 pounds of food and 5 pieces of clothing and one containing 100 pounds of food and 2 pieces of clothing. These two bundles are represented by two points, K and L , in Figure 3.1. In addition, other market bundles, each of which is represented by a point in Figure 3.1, are just as desirable to Ms. Popovich. If we connect all these points, we derive a curve that represents equally desirable bundles to Ms. Popovich. Figure 3.1 maps these bundles as points on curve I_1 in Figure 3.1. Curve I_1 is an indifference curve.

We need to understand three things when modeling consumer indifference curves:

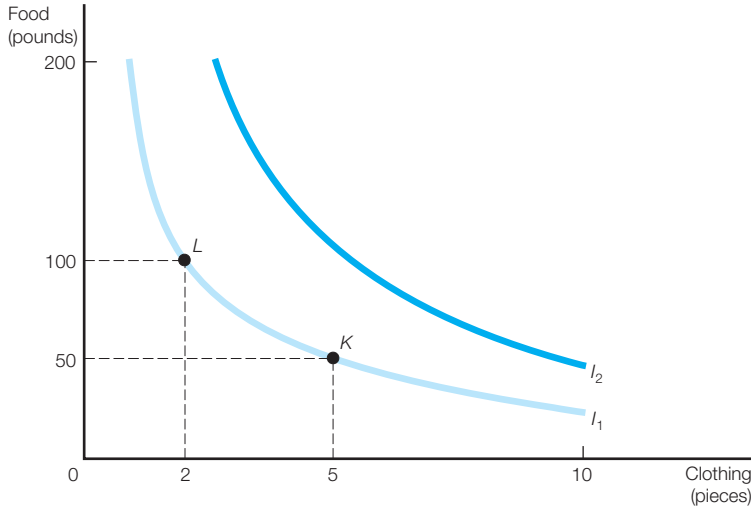
1. *A consumer has many indifference curves.* If Ms. Popovich is indifferent among all the market bundles represented by points on I_2 in Figure 3.1, I_2 is

Indifference curve Contains points representing market bundles among which the consumer is indifferent.

FIGURE 3.1

Two of Ms. Popovich's Indifference Curves

The curves I_1 and I_2 are two of Ms. Popovich's indifference curves. Each shows market bundles that are equally desirable to Ms. Popovich.



another of her indifference curves. Moreover, one thing is certain. She prefers any I_2 bundle to any one on I_1 , because I_2 has bundles with as much clothing as and more food than (or as much food as and more clothing than) bundles on I_1 . Implicitly we assume that consumers are sometimes insatiable. (Of course consumers sometimes become so satiated with a product that they prefer less of it to more, but we assume for simplicity that this is not the case here.) Consequently, market bundles on higher indifference curves like I_2 are preferred to bundles on lower indifference curves like I_1 .

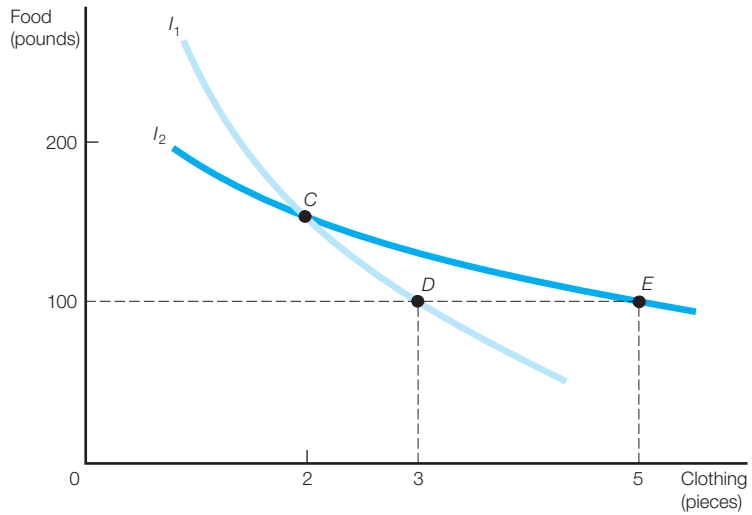
2. *Every indifference curve must slope downward and to the right, so long as the consumer prefers more of each commodity to less.* If one market bundle on an indifference curve has more of one product than a second bundle, it must have less of the other product than the second bundle. This is true so long as more of each product is preferred.

3. *Indifference curves cannot intersect.* If they did, this would contradict the assumption that more of a product is preferred. For example, suppose that I_1 and I_2 in Figure 3.2 are two intersecting indifference curves. If this is the case, the market bundle at point D is equivalent to the one represented by point C because both are on indifference curve I_1 . Moreover, the market bundle represented by point E is equivalent in the eyes of the consumer to the one represented by point C because

FIGURE 3.2

Intersecting Indifference Curves: A Contradiction

Indifference curves cannot intersect. If they did, the consumer would be indifferent between D and C because both are on indifference curve I_1 , and between E and C because both are on indifference curve I_2 . But this implies that he or she must be indifferent between D and E , which is impossible because E contains the same amount of food and two more pieces of clothing than D , and we are assuming that more of a commodity is preferred to less.



both are on indifference curve I_2 . And this means the market bundle represented by point E is equivalent to the one represented by point D . But this is impossible because the bundle at E contains the same amount of food and two more pieces of clothing than bundle D . If we assume that more of a product is preferred, the bundle at E must be preferred to the bundle at D .

THE MARGINAL RATE OF SUBSTITUTION

Some consumers place a high value on obtaining an extra unit of a product; others place a low value on obtaining it. If managers are to understand consumer choice, it is useful to measure the relative importance a consumer places on acquiring an additional unit of a particular product. We measure this using what is called the marginal rate of substitution.

The **marginal rate of substitution** of product X for product Y is defined as the number of units of product Y that must be given up if the consumer, after receiving an extra unit of product X , is to maintain a constant level of satisfaction. Obviously the more units of product Y the consumer is willing to give up to get an

Marginal rate of substitution The number of units of product Y that must be given up if the consumer, after receiving an extra unit of product X , is to maintain a constant level of satisfaction.

FIGURE 3.3

Indifference Curves of Consumers with High and Low Marginal Rates of Substitution of Performance for Stylishness

The left panel shows the indifference curves of consumers who are willing to trade a lot of stylishness for a little extra performance. The right panel shows the indifference curves of consumers who are willing to trade a lot of performance for a little extra stylishness.



extra unit of X , the more important product X is (relative to Y) to the consumer. To estimate the marginal rate of substitution, we multiply the slope of the consumer's indifference curve by -1 . This gives us the number of units of product Y the consumer is willing to give up for an extra unit of product X .

To illustrate, consider consumer preferences for attributes of automobiles. Two key attributes are stylishness and performance (for example, speed, gasoline mileage, and handling). Some consumers are willing to trade a lot of stylishness for a little extra performance. This behavior is shown in the left panel of Figure 3.3. The indifference curves here are steep. The marginal rate of substitution of performance for stylishness is relatively high because the slope of the indifference curves (times -1) is relatively large. Other consumers are willing to trade a lot of performance for a little extra stylishness. For these consumers the indifference curves in Figure 3.3 are relatively flat, as in the right panel of Figure 3.3. The marginal rate of substitution of performance for stylishness is relatively low because the slope of the indifference curves (times -1) is relatively small.

THE CONCEPT OF UTILITY

The consumer's indifference curves represent his or her tastes and preferences. Given all the indifference curves of a particular consumer, we attach a number, called a utility, to each of the available market bundles. **Utility** indicates the level of enjoyment or preference attached by a consumer to a particular market bundle.

Utility Indicates the level of enjoyment or preference attached by a consumer to a particular market bundle.

More specifically, it summarizes the preference ranking of market bundles; the higher the utility assigned to a bundle, the higher the level of satisfaction the consumer realizes from it. Because all market bundles on a given indifference curve yield the same amount of satisfaction, they all have the same utility. Market bundles on higher indifference curves have higher utilities than those on lower indifference curves.

When we assign utilities to market bundles, it tells us which bundles the consumer prefers. If the utility attached to one bundle is higher than that attached to another, the consumer prefers the first over the second. If the utility attached to the first bundle is lower than the second, he or she prefers the second over the first. If the utility attached to the first market bundle equals the second, he or she is indifferent between the two market bundles.

How do we rationally assign these utilities? Assume our consumer prefers market bundle R to bundle S , and bundle S to bundle T . The utility assigned to bundle R must be higher than that assigned to bundle S , while the utility assigned to bundle S must be higher than that assigned to bundle T . Any set of numbers conforming to these requirements is an adequate measure of utility. Therefore, the utility of market bundles R , S , and T may be 30, 20, and 10 or 6, 5, and 4, respectively. All that counts is that the utility of market bundle R is higher than that of bundle S , which in turn should be higher than that of bundle T . Put differently, both sets of utilities provide a correct ordering or ranking of market bundles in terms of levels of consumer satisfaction.

Indifference curves are also known as iso-utility curves. We can measure the slope described above as

$$-\Delta f/\Delta c = -(\Delta U/\Delta c)/(\Delta U/\Delta f) = -MU_c/MU_f$$

where MU_f is the marginal utility of food, that is, the increase in Ms. Popovich's utility if she obtains one more unit of food (holding the amount of clothing she possesses constant), and where MU_c is the marginal utility of clothing, that is, the increase in Ms. Popovich's utility if she obtains one more unit of clothing (holding the amount of food she possesses constant). Thus, Ms. Popovich's marginal rate of substitution is equal to the ratio of her marginal utility of clothing to her marginal utility of food. These marginal utilities are precisely what we were talking about earlier in this section when we referenced the value a consumer placed on obtaining an extra unit of a product.

THE BUDGET LINE

Consumers wish to maximize their utility, which means they want to consume bundles from the highest possible indifference curve. But whether a particular indifference curve is attainable depends on a consumer's income and product prices. To

make things concrete, we return to our consumer, Jennifer Popovich. Suppose her total income is \$600 per week, and she spends it on only food and clothing.

How much of each product Ms. Popovich can buy depends on the prices of food and clothing. Suppose a pound of food costs \$3 and a piece of clothing costs \$60. Then, if she spent all her income on food, she could buy 200 pounds of food per week. On the other hand, if she spent all her income on clothing, she could buy 10 pieces of clothing per week. Or she could, if she wished, buy some food and some clothing. There are many combinations of food and clothing she could buy, and each such combination can be represented by a point on the line in Figure 3.4. This line is called her budget line. A consumer's **budget line** shows the market bundles that he or she can purchase, given the consumer's income and prevailing market prices.

To obtain the equation for Jennifer Popovich's budget line, note that

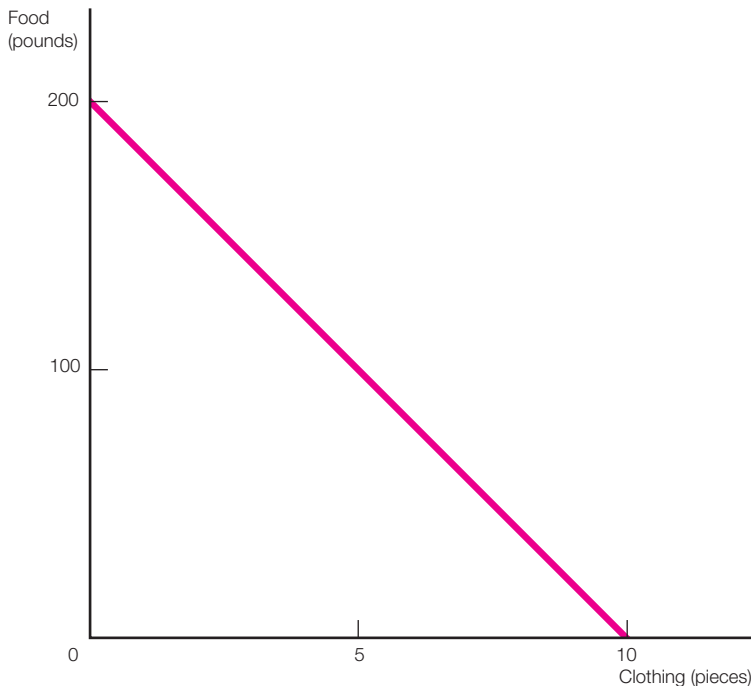
$$YP_f + XP_c = I \quad (3.1)$$

Budget line Shows the market bundles that the consumer can purchase, given the consumer's income and prevailing market prices.

FIGURE 3.4

Ms. Popovich's Budget Line

The consumer's budget line shows the market bundles that can be purchased, given the consumer's income and prevailing commodity prices. This budget line assumes that Ms. Popovich's income is \$600 per week, that the price of a pound of food is \$3, and that the price of a piece of clothing is \$60.



where Y is the amount of food she buys, X is the amount of clothing she buys, P_f is the price of food, P_c is the price of clothing, and I is her income. The left side of equation (3.1) equals the total amount she spends on food and clothing; what equation (3.1) says is that this amount must equal her income. For simplicity, we assume she saves nothing. (This assumption can be relaxed.) Solving equation (3.1) for Y , we obtain

$$Y = \frac{I}{P_f} - \frac{P_c}{P_f} X \quad (3.2)$$

which is the equation for her budget line.

A shift occurs in a consumer's budget line if changes occur in the consumer's income or product prices. In particular, an increase in income raises the budget line, whereas a decrease in income causes the budget line to fall (parallel to the original line because a change in I does not affect the slope). This is illustrated in Figure 3.5, which shows Ms. Popovich's budget lines at incomes of \$300, \$600, and \$900 per week. Her budget line moves upward as her income rises.

Also, the prices of products affect the budget line. A decrease in a product's price causes the budget line to intersect this product's axis at a point farther from

FIGURE 3.5

Ms. Popovich's Budget Lines at Incomes of \$300, \$600, and \$900 per Week

The higher the consumer's income, the higher is the budget line. Holding commodity prices constant, the budget line's slope remains constant.

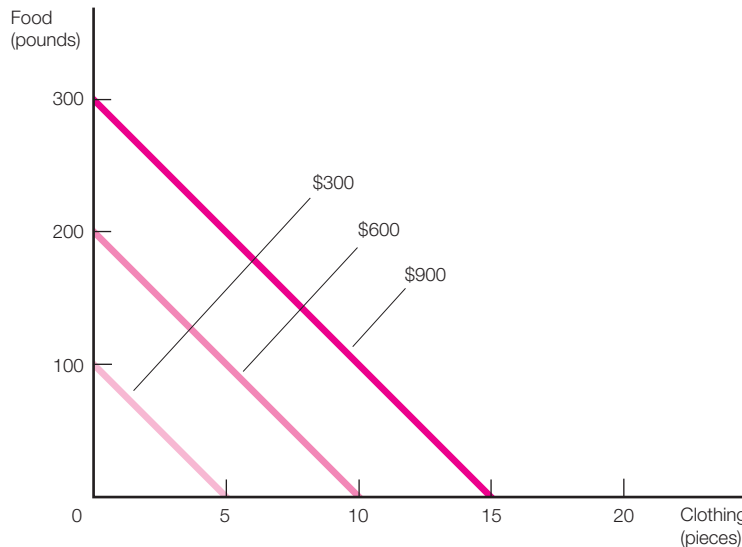
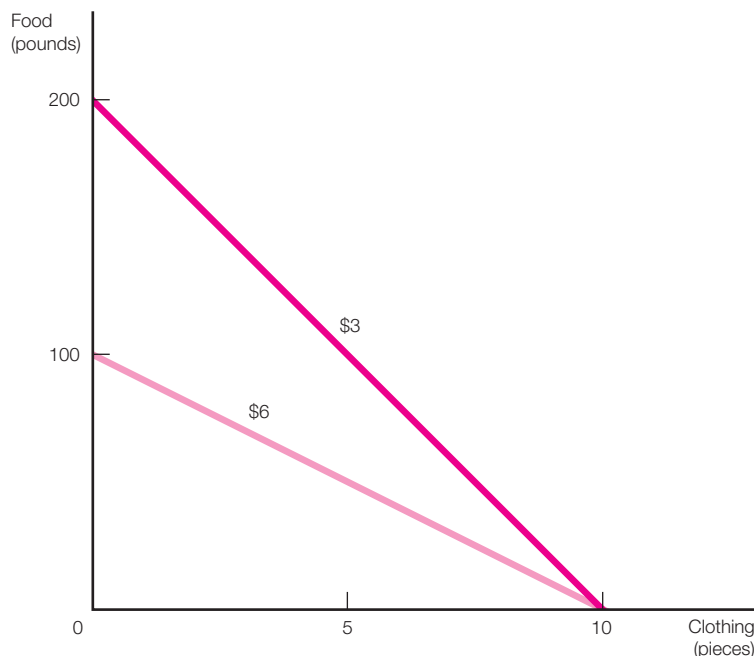


FIGURE 3.6

Ms. Popovich's Budget Line at Food Prices of \$3 and \$6 per Pound

Holding constant Ms. Popovich's income at \$600 per week and the price of a piece of clothing at \$60, the budget line intersects the vertical axis farther from the origin when the price of food is \$3 than when it is \$6.



the origin. Figure 3.6 shows Ms. Popovich's budget line when the price of a pound of food is \$3 and when it is \$6. You can see the budget line meets the vertical, or food, axis farther from the origin when the price of food is \$3 per pound. This is because the change in the price of food alters the slope of the budget line, which equals $-P_c/P_f$ (as shown in equation (3.2)).

THE EQUILIBRIUM MARKET BUNDLE

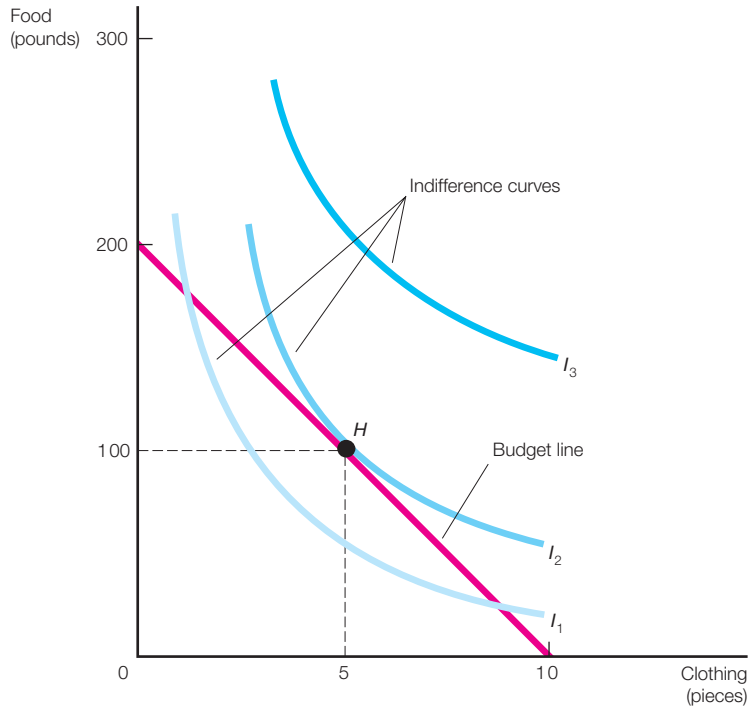
Given a consumer's indifference curves and budget line, we can determine the consumer's **equilibrium market bundle**—the market bundle that, among all the items the consumer can purchase, yields the maximum utility. The first step is to combine the indifference curves with the budget line on the same graph. Figure 3.7 brings together Ms. Popovich's indifference curves (from Figure 3.1) and her budget line (from Figure 3.4). On the basis of the information assembled in Figure 3.7, it is a simple matter to determine her equilibrium market bundle. Her indifference

Equilibrium market bundle The market bundle that, among all the items the consumer can purchase, yields the maximum utility.

FIGURE 3.7

Equilibrium Market Bundle

Ms. Popovich's equilibrium market bundle is at point H , containing 100 pounds of food and 5 pieces of clothing. This is the point on her budget line that is on the highest indifference curve she can attain, I_2 .



curves show what she wants: Specifically, she wants to attain the highest possible indifference curve. Therefore, she would rather be on indifference curve I_2 than on indifference curve I_1 and on indifference curve I_3 than on indifference curve I_2 . But she cannot choose any market bundle she likes. The budget line shows which market bundles her income and product prices permit her to buy. Consequently she must choose a bundle on her budget line.

Clearly the consumer's choice boils down to choosing the market bundle on the budget line that is on the highest indifference curve. This is the equilibrium market bundle. For example, Ms. Popovich's equilibrium market bundle is at point H in Figure 3.7; it consists of 100 pounds of food and 5 pieces of clothing per week. This is her equilibrium market bundle because any other bundle on the budget line is on a lower indifference curve than point H . But will Ms. Popovich choose this bundle? It may take some time for her to realize this is the best market bundle under the circumstances, but eventually we should expect her to purchase this bundle.

STRATEGY SESSION: The Effect of a Time Constraint on Consumer Behavior

For consumers, time can be as important as money. For example, suppose Mildred Evans, an avid sports fan who goes regularly to baseball and football games, decides that she can devote no more than 24 hours per month to attending such games and that she can spend no more than \$120 per month on baseball and football tickets. She lives much closer to the local baseball stadium than to the nearest football stadium, so it takes her 4 hours to see a baseball game but 6 hours to see a football game. The price of each baseball ticket is \$10, and the price of each football ticket is \$40.

Let B be the number of baseball games and F be the number of football games she attends per month. If she spends a total of \$120 per month on tickets,

$$40F + 10B = 120 \quad (3.3)$$

Why? Because $40F$ is the amount spent on football tickets, and $10B$ is the amount spent on baseball tickets, so $40F + 10B$ is the total amount spent per month on baseball and football tickets, which must equal \$120. From equation (3.3), it follows that

$$F = 3 - B/4 \quad (3.4)$$

This is the equation for the budget line, plotted in the following graph.

But this ignores the time constraint. If she spends a total of 24 hours per month at baseball and football games,

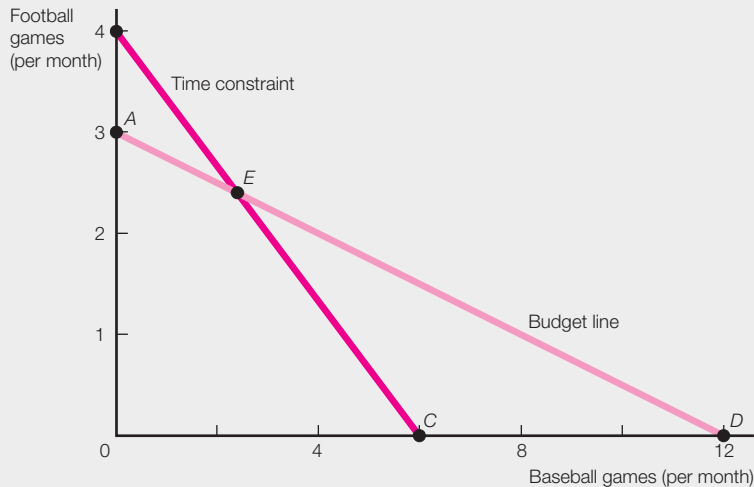
$$6F + 4B = 24 \quad (3.5)$$

Why? Because $6F$ equals the number of hours spent at football games and $4B$ equals the number of hours spent at baseball games, so $6F + 4B$ equals the total number of hours spent at baseball and football games, which must equal 24. From equation (3.5), it follows that

$$F = 4 - 2B/3 \quad (3.6)$$

This is the equation for the time constraint, plotted in the graph.

To keep within both the time and expenditure constraints, Mildred must pick a market bundle on line segment AE or line segment EC in the graph. Note that the time constraint cuts down on the number of feasible market bundles. Given that she wants to devote only 24 hours per month to attending baseball and football games, she must be content with market bundles along line segment EC rather than line segment ED , which would be available if there were no time constraint.



MAXIMIZING UTILITY: A CLOSER LOOK

Let us look more closely at the equilibrium market bundle at point H , the one Ms. Popovich chooses. Clearly this bundle is at the point where the budget line is tangent to an indifference curve. Because the slope of the indifference curve equals -1 times the marginal rate of substitution of clothing for food (see page 69) and the slope of the budget line is $-P_c/P_f$ (see page 73), it follows that Ms. Popovich, if she maximizes utility, chooses in equilibrium to allocate her income between food and clothing so that

$$MRS = P_c/P_f \quad (3.7)$$

where MRS is the marginal rate of substitution of clothing for food.

To understand what this means, recall that the marginal rate of substitution is the rate at which the consumer is willing to substitute clothing for food, holding her total level of satisfaction constant. Hence if the marginal rate of substitution is 4, the consumer is willing to give up 4 pounds of food to obtain 1 more piece of clothing. On the other hand, the price ratio, P_c/P_f , is the rate at which the consumer is able to substitute clothing for food. So if P_c/P_f is 4, the consumer *must* give up 4 pounds of food to obtain one more piece of clothing.

What equation (3.7) is saying is this: The rate at which the consumer is willing to substitute clothing for food (holding satisfaction constant) must equal the rate at which he or she is able to substitute clothing for food. Otherwise, it is always possible to find another market bundle that increases the consumer's satisfaction.

To illustrate, suppose Ms. Popovich chooses a market bundle for which the marginal rate of substitution of clothing for food is 4. Suppose the price ratio, P_c/P_f , is 3. If this is the case, Ms. Popovich can obtain an extra piece of clothing if she buys 3 fewer pounds of food because the price ratio is 3. But an extra piece of clothing is worth 4 pounds of food to Ms. Popovich because the marginal rate of substitution is 4. Therefore, she can increase her satisfaction by substituting clothing for food—and this will continue to be the case so long as the marginal rate of substitution exceeds the price ratio. Conversely, if the marginal rate of substitution is less than the price ratio, Ms. Popovich can increase her satisfaction by substituting food for clothing. Only when the marginal rate of substitution equals the price ratio does her market bundle maximize her utility.

CORNER SOLUTIONS

Although in our example Ms. Popovich chooses the market bundle where the budget line is tangent to an indifference curve (the market basket at point H in Figure 3.7), this is not always true. A consumer may consume none of some

STRATEGY SESSION: A Manager's Trade-Off Between Output and Profit

In the corporate world, managers are often asked to choose between multiple goals. We can use indifference curves to estimate this behavior. Let's say a manager is driven by two goals: (1) She wants to make a large profit, and (2) she also wants to be noticed. And small firms don't tend to get noticed while big firms do.

For simplicity, assume she is a monopolist, so the market demand curve is her firm demand curve. The demand curve is

$$P = a - bQ$$

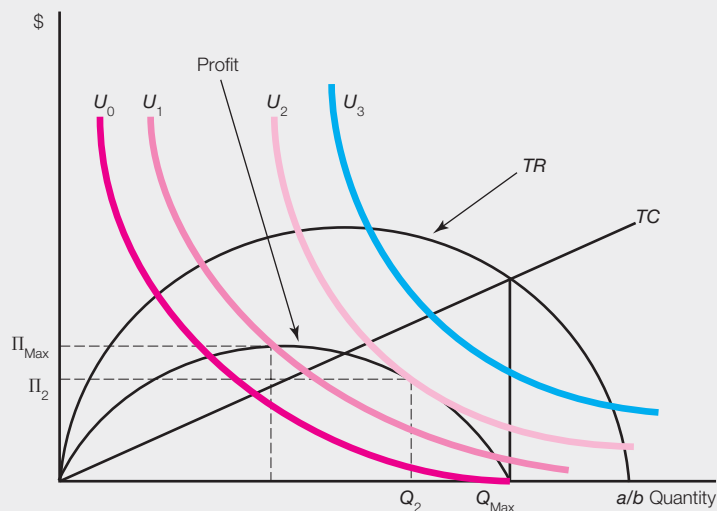
Hence $P = a$ when $Q = 0$ and $Q = a/b$ when P is 0. Total revenues are 0 when $P = 0$ and 0 when $Q = 0$. She can maximize revenues at a price where $\eta = -1$. If it costs a constant k to make a unit of product, total costs are $TC = kQ$. The figure below visualizes total revenue, total cost, and profit (the difference between the two).

The manager's utility curves have the usual shape: Utility increases if output is held constant and profit increases, and utility increases if profit is held constant and output increases. Of course she is

happiest if both output and profit increase. The profit curve now is her constraint curve. She will maximize utility with profit of Π_2 and output of Q_2 generating utility of U_2 . Note that she does not maximize profit (which gives utility of only U_1) or quantity (because that gives her utility of only U_0 —we're constraining her output size objectives to profitable output).

Now let's say our manager works for a large, publicly held corporation. You are a shareholder of that company, and you'd prefer to maximize profit. There is ample evidence to support your preference because higher profit is strongly associated with higher stock prices. We have just shown that the manager does not want to maximize profit. What can the shareholder do to revise the behavior of our manager so she also wants to maximize profit?

Such issues are faced by shareholders of most publicly held corporations. We will discuss these behaviors in Chapter 15 when we discuss the principal-agent issue. But as we can see from the figure, if we tie the manager's compensation to firm profit, our manager will care more about maximizing profit.



products because the cost does not justify the level of satisfaction they receive. For example, though many people have the income to afford a small amount of Beluga caviar, they do not purchase it because the cost is greater than the pleasure they receive from eating fish eggs.

Figure 3.8 shows this situation graphically. For simplicity, we assume the consumer can choose only between Beluga caviar and pizza. Given the position of these indifference curves, the consumer maximizes utility with bundle W , which contains all pizza and no Beluga caviar. This market bundle maximizes utility because it is on a higher indifference curve than any other bundle on the budget line. It is a **corner solution** in which the budget line touches the highest achievable indifference curve along an axis (in this case the vertical axis).

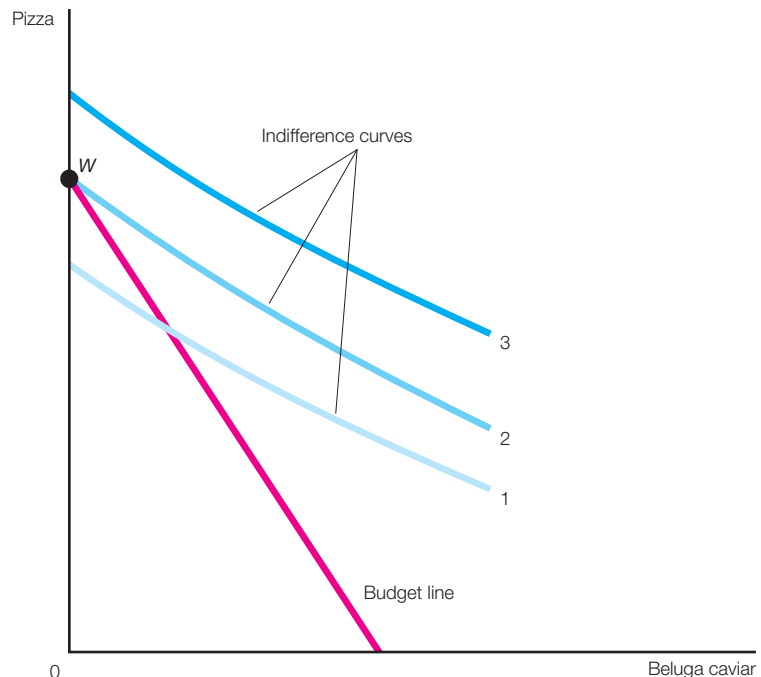
We previously showed that if the consumer purchases some of both goods to maximize utility, the marginal rate of substitution is equal to the price ratio. But if the consumer maximizes utility with a corner solution, this is not the case.

Corner solution When the budget line touches the highest achievable indifference curve along an axis.

FIGURE 3.8

Corner Solution

The market bundle that maximizes your utility is W , which lies on the vertical axis.



HOW MANAGERS CAN STRATEGICALLY INFLUENCE CONSUMER CHOICES

We just saw how Ms. Popovich's purchase decisions are influenced by her preferences, her income, and product prices. Although managers can do little to change consumer incomes, they can influence preferences and the effect of a budget constraint. For example, advertising is a direct action to influence preferences; lowering prices may induce a consumer to purchase products. So managers can influence budget constraints with their pricing policies. We have just portrayed the budget constraint as linear to present the basic theory of consumer choice; but in the real world managers price in ways that are consciously designed to make the budget constraint nonlinear and to influence consumer choice.

A recent coupon offer by a leading grocery chain (Albertson's) offered its customers \$18 off their grocery bill if they spent at least \$180 in one store visit. Let's view the impact of such an offer on Ms. Popovich. Suppose her income is \$200. She can consume groceries at \$1 per unit or clothing at \$1 per unit. Thus, before the coupon, her budget constraint and purchase decision (she purchases C_0 units of clothing, G_0 units of groceries, and attains utility level U_0) appears in Figure 3.9.

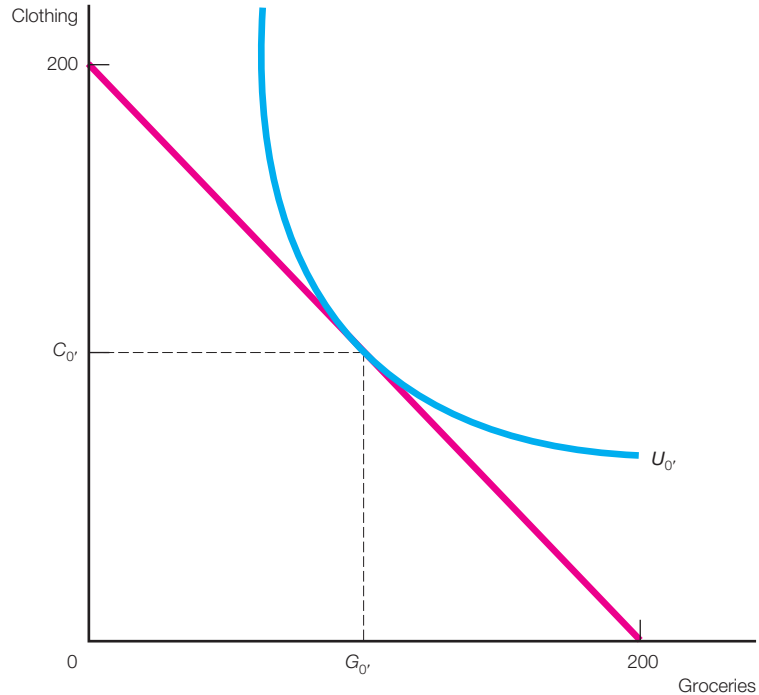
By offering a coupon, Albertson's managers in effect shift Ms. Popovich's budget constraint outward parallel to her old one. When her grocery bill reaches \$180 (so she consumes \$20 of clothing), she receives \$18 back (or pays \$162 to the grocery store) and now has the potential of spending \$38 on clothing (if she spends all her savings on clothing). Alternatively, she could spend \$218 on groceries (if she took her \$18 in savings and spent it on more groceries). Her new budget constraint and purchase choices appear in Figure 3.10.

The coupon budget constraint looks like her no-coupon budget constraint until she spends at least \$180 on groceries. Then her budget constraint shifts upward to the right by an additional \$18. In effect, through their pricing decisions, Albertson's managers have shifted Ms. Popovich's budget constraint. If Ms. Popovich's indifference curves resemble the dashed line in Figure 3.10, the coupon does not affect her purchase behavior (her utility stays at U_0 , and her purchases are C_0 and G_0). Nothing was gained by Albertson's managers, but the cost was trivial (the printing of some coupons). However, had Ms. Popovich's initial indifference curve been the solid U_0 , the coupon would enable her to increase her utility to U_1 . The coupon was a good deal for Ms. Popovich. Was it a good strategy by Albertson's managers? Only if $G_1 - G_0 > 18$; that is, only if Ms. Popovich spent over \$18 more on groceries than she did without the coupon. Presumably the managers felt that most consumers would spend more than \$18 extra on groceries as a result of the coupon. So we see that managers have a range of strategies they can use to change the purchasing decisions of consumers in addition to advertising and changing prices.

FIGURE 3.9

Ms. Popovich's Utility-Maximizing Purchase of Clothing and Groceries before Receiving the Coupon

With a budget of \$200 and the price of a unit of clothing and a unit of groceries each \$1, Ms. Popovich rationally chooses to consume C_{0^*} units of clothing and G_{0^*} units of groceries, attaining a utility of U_{0^*} .

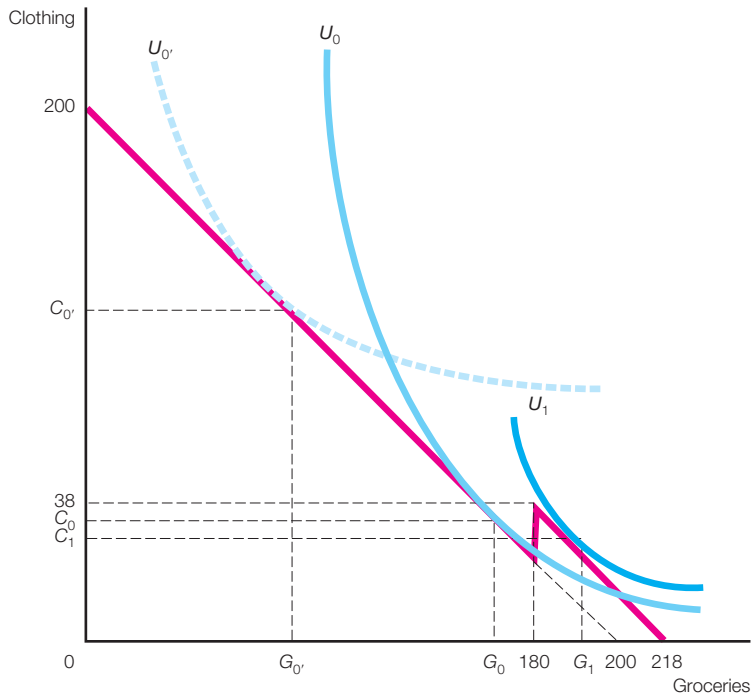


Another strategy for influencing the budget constraint of consumers is quantity discounts. Ms. Popovich pays a visit to Dunking Donuts. A single donut costs \$.50, a half dozen costs \$2, and a dozen costs \$3. Suppose Ms. Popovich has \$4 to spend on donuts and all other goods. The price of a unit of all other goods is \$1. Thus if Ms. Popovich buys one donut, she will have \$3.50 to spend on all other goods. If she buys two donuts, she will have \$3 to spend on all other goods. If Ms. Popovich buys four donuts separately, she will spend \$2 on donuts and \$2 on all other goods—but she could buy a half dozen for \$2. If more is indeed better (that is, if the marginal utility of donuts is positive), she should buy the half dozen rather than four or five donuts. If Ms. Popovich wants seven donuts, she should buy the half dozen for \$2 and one donut separately for \$0.50; this will leave her with \$1.50 for all other goods. Should she wish to buy eight donuts, she should buy the dozen for \$3. She could spend \$2 for a half dozen and \$0.50 each for donuts

FIGURE 3.10

Ms. Popovich's Utility-Maximizing Purchase of Clothing and Groceries after Receiving the Coupon

Part of Ms. Popovich's budget line shifts outward, parallel to her budget line without the coupon. Depending on the shape of her indifference curves, the coupon may or may not influence her purchases of clothing and groceries. If her indifference curves are similar to the dashed line here, her behavior is uninfluenced; but if her indifference curves are like the solid line, the coupon increases her utility (and perhaps firm revenue).



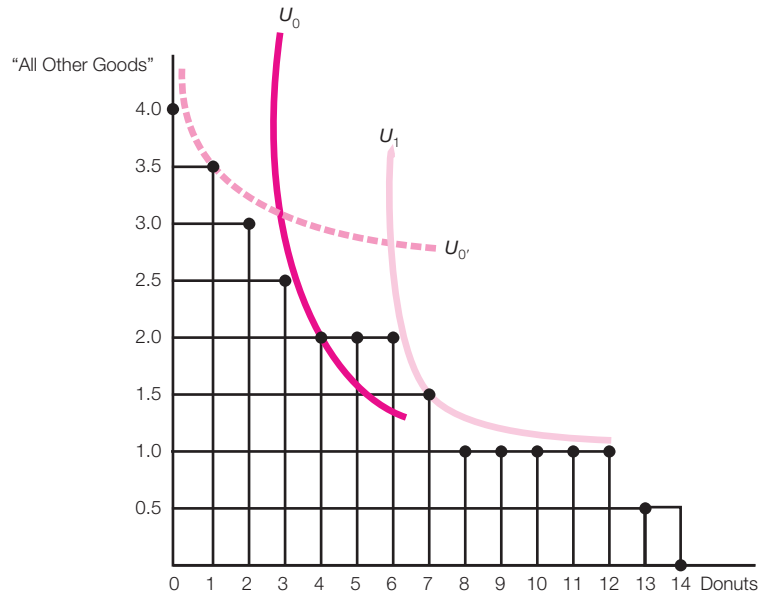
seven and eight, or she could get a dozen for the same \$3. Thus she should never buy 8, 9, 10, or 11 donuts. Her budget constraint will appear as the step function in Figure 3.11.

In effect, what is the influence of the pricing strategy? Selling a half dozen donuts for \$2 reduces the price per donut to \$0.33, and selling a dozen donuts for \$3 reduces the price per donut to \$0.25. Why should the managers offer this price decrease? Perhaps because of the diminishing marginal utility of donuts. Consumers are willing to pay a lot for the first donut. But donuts are filling (and fattening). As the marginal utility of donuts decreases, the Dunking managers must lower the price to entice buyers.

FIGURE 3.11

Ms. Popovich's Utility-Maximizing Choices for Donuts and All Other Goods

The quantity discount for donuts creates a step function budget constraint and means that Ms. Popovich will never purchase 4, 5, 8, 9, 10, or 11 donuts. The quantity discount doesn't change her behavior if her indifference curves resemble the dashed line but raises her utility if her indifference curves resemble the solid line.



If the price of donuts was always \$0.50/donut, Ms. Popovich's budget constraint intercept on the donuts axis would occur at 8 donuts. With the quantity discount, her budget constraint intercept is at 14 donuts. If Ms. Popovich's indifference curve is the dashed one, the quantity discount strategy will not change her behavior; but if it is the solid one, she is induced to move from buying 4 donuts to 7 while increasing her utility from U_0 to U_1 . Both managers and consumers can benefit from this strategy.

You face this situation anytime you are offered a quantity discount. Remember this the next time you encounter one and see whether the pricing strategy changes your behavior.

We can also use shifts in the budget constraint to explain why most individuals prefer cash to a specific gift (unless they requested the gift). At every holiday season, consumers decide what gifts to buy for friends and family. Suppose Ms. Popovich is already maximizing her utility subject to her budget constraint (I_0) with A_0 units of all other goods and G_0 units of the gift good and receiving U_0 units of satisfaction. Let the unit price of all other goods, P_A , be 1. Suppose

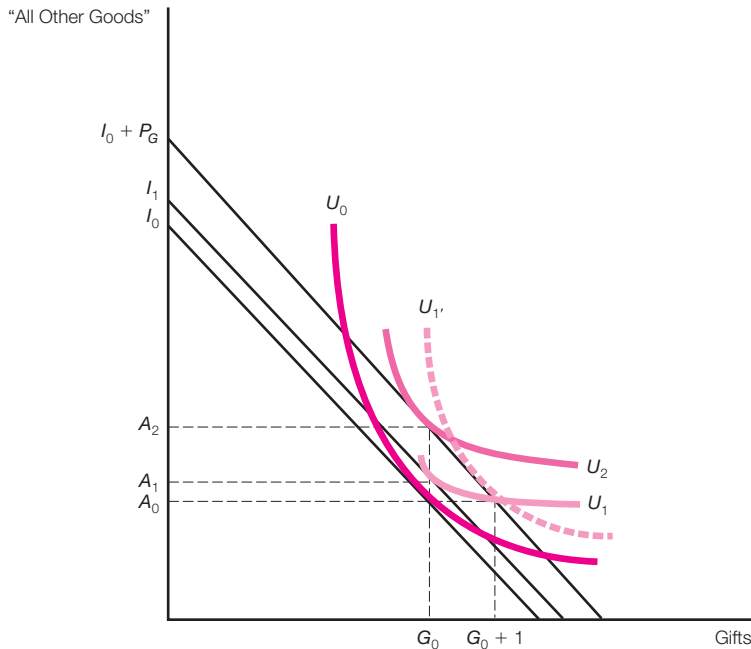
Ms. Popovich's well-meaning mother-in-law gives her another unit of the gift item for the holidays. Assuming the marginal utility of the gift item is positive, Ms. Popovich's utility has increased. Gift giving has been a success as far as her mother-in-law is concerned. But how about for Ms. Popovich? Consider the situation depicted in Figure 3.12.

After receiving the gift unit from her mother-in-law, if Ms. Popovich's high-level utility resembles the dashed line, she will have A_0 units of all other goods and $G_0 + 1$ units of the gift good and will maximize her utility at level $U_{1'}$. By spending P_G on a gift unit, the mother-in-law got it right on. The mother-in-law is happy that Ms. Popovich appreciates the gift, and Ms. Popovich is happy because her utility is increased. You will be lucky if all your gift giving turns out this way. Unfortunately, in many cases the most likely scenario is depicted by the solid higher-level indifference curves (U_1 and U_2). We can state two things. If the

FIGURE 3.12

Ms. Popovich's Utility and Consumption of All Other Goods and Gift Goods under Various Scenarios of Gift Receiving

A gift will, at best, give Ms. Popovich the exact consumption of the gift good and all other goods had Ms. Popovich made the choice herself. But most likely a smaller cash gift from the gift giver will yield Ms. Popovich the same utility as the gift; or a gift of cash to Ms. Popovich equal to what the gift giver spent will yield a higher level of utility than did the gift.



mother-in-law's objective is to get Ms. Popovich to a level of satisfaction U_1 , she could do so with a smaller expenditure ($I_1 < I_0 + P_G$), giving Ms. Popovich no gift units and instead $A_1 - A_0$ units of all other goods. Ms. Popovich will get the same level of utility as with the gift if she had A_1 units of all other goods and G_0 gift units. In general, from a utility standpoint, gift giving is more expensive than it has to be. A second point is that if the mother-in-law gives Ms. Popovich the cash she spent on her gift (P_G), Ms. Popovich's budget line will shift outward to $I_0 + P_G$ and Ms. Popovich will maximize her utility by consuming A_2 units of all other goods and G_0 units of the gift good. That is, a cash gift equal to the unit cost of the gift good will increase Ms. Popovich's utility to U_2 , which is higher than the U_1 level attained with the gift. Thus the mother-in-law could spend less and give Ms. Popovich the same level of satisfaction by buying Ms. Popovich more units of all other goods (that is, by giving cash) rather than the gift she bought her.

Our colleague Joel Waldfogel has received much "Grinch" publicity since his article about this deadweight loss of Christmas was published in 1993.¹ In fact, every year around Christmas he gets calls from newspapers looking for seasonal stories about Christmas. If you look at the proliferation of gift cards in recent years, you might conclude that his concept is getting through. However, gift cards tie the recipient to a particular store: A gift card to an apparel store may not be good if you really want electronics. General gift cards, like those from American Express, are less constraining but aren't universally accepted; cash is accepted in all brick-and-mortar stores. We should point out, though, that our analysis does not account for sentimental value. You may hate the gift that your grandmother gives you but be happy knowing that she loves you enough to select and purchase it.

DERIVING THE INDIVIDUAL DEMAND CURVE

A consumer's demand curve shows how much product a person will purchase at various prices (when other prices, preferences, and income are held constant). It reveals the inner desires of purchase behavior. Let us return to Ms. Popovich.

Ms. Popovich can choose between two products: food and clothing. Her weekly income is \$600, and the price of clothing is \$60 per piece. Ms. Popovich's budget is budget line 1 in Figure 3.13 when the price of food is \$3 per pound. As we saw in Figure 3.7, she will buy 100 pounds of food per week.

How will she purchase when the price of food increases to \$6 per pound? If her income and the price of clothing remain constant, her budget is budget line 2 in Figure 3.14. She attains her greatest utility by reaching her highest indifference curve, I_1 . She chooses the bundle at point K , which contains 50 pounds of food per week. If the price of food is \$6 per pound, she will make a weekly purchase of 50 pounds of food per week.

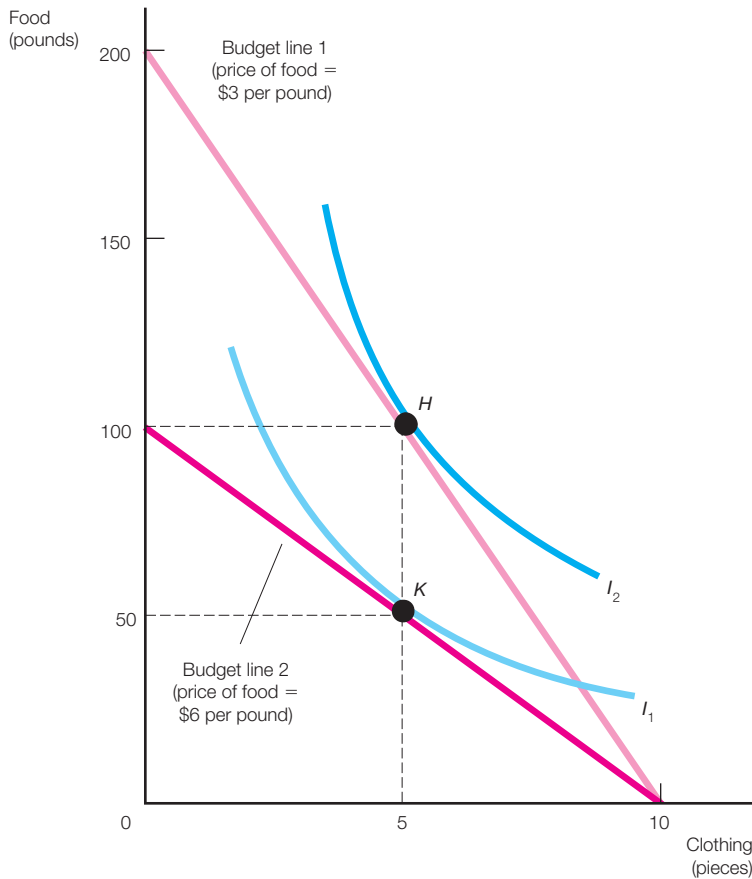
We have derived two points on Ms. Popovich's demand curve for food, those corresponding to food prices of \$3 and \$6 per pound. Figure 3.14 shows these

1. J. Waldfogel, "The Deadweight Loss of Christmas," *American Economic Review* Vol. 83(5) (December 1993), pp. 1328–1336.

FIGURE 3.13

Effect of a Change in Price on Ms. Popovich's Equilibrium Market Bundle

If the price of a pound of food is \$3, Ms. Popovich's budget line is such that her equilibrium market bundle is at point *H*, where she buys 100 pounds of food per week. If the price of a pound of food is \$6, Ms. Popovich's budget line is such that her equilibrium market bundle is at point *K*, where she buys 50 pounds of food per week.



two points, *U* and *V*. To obtain more points on her demand curve, all we do is assume a particular price of food, construct the budget line corresponding to this price (holding her income and the price of clothing constant), and find the market bundle that is on her highest indifference curve. Plotting the level of food in this bundle relative to the assumed price of food, we obtain a new point on her demand curve for food. Connecting all these points, we estimate her complete demand curve for food, in Figure 3.14. (In our scenario, the level of clothing consumed remains constant. This does not have to be the case.)

PROBLEM SOLVED: Do I Stay or Do I Go? Use of Indifference Curves

Ms. Popovich is a self-employed business owner who is also raising a family. Her business is growing, and she finds it could literally demand 18–20 hours per day. Of course she feels the same about her children. What is she to do? We call the time spent engaged in business *work* and time spent not working *leisure*.^a Unlike money budget constraints, which vary with an individual's income, the time constraint is a great equalizer: There are 24 hours each day whether you are rich or poor. Thus Ms. Popovich is directed by a time budget constraint of $H_W + H_L = 24$, where H_W is the hours worked and H_L is the hours spent in leisure pursuits. We express hours worked as $H_W = 24 - H_L$.

In every hour working, Ms. Popovich receives a wage of W . We use that knowledge to estimate a time budget constraint with income. Her utility (level of satisfaction) is a function of income (I) and leisure (L): $U = U(I, L)$. As in a purchase decision, she must choose between the two. For a given level of income, she'd be happier with more leisure, and for a given time in leisure, she'd be happier with more income. Her indifference curves are downward-sloping and convex to the origin (the "normal" shape).

We estimate Ms. Popovich's income as the hourly wage (W) times the hours worked

$$W \times (24 - H_L) = 24W - W \times H_L$$

We presuppose that time is money and show why with a budget constraint. If she chooses all work and no play, $H_W = 24$ and $H_L = 0$, her income is $24W$. If she chooses all play and no work, $H_W = 0$ and $H_L = 24$, her income is 0. Her utility-maximizing behavior is shown in the figure.

The slope of the time constraint is $-W = -24W/24$; that is, the market will pay her W to work an hour (or alternatively, she forgoes W for every hour of leisure). The slope of her indifference curve is

$$\begin{aligned} \frac{-\Delta I}{\Delta L} &= \left(\frac{-\Delta I}{\Delta L} \right) \left(\frac{\Delta U}{\Delta U} \right) \\ &= - \left(\frac{\Delta U}{\Delta L} \right) / \left(\frac{\Delta U}{\Delta I} \right) = - \frac{MU_L}{MU_I} \end{aligned}$$

Or more elegantly,

$$- \frac{MU_L}{MU_I} = - \left(\frac{dU}{dL} \right) / \left(\frac{dU}{dI} \right)$$

We estimate that Ms. Popovich will act as if setting $MU_L/MU_I = W$, that is, the slope of the indifference curve, is equal to the slope of the constraint. That is, she works up to the point where her trade-off of leisure for income equals the wage rate. By changing the wage rate (for instance, from W to a greater rate W'), we can estimate how Ms. Popovich changes her demand for leisure (in this case choosing less of it, $H_L' < H_L$, and preferring more income). We can change the wage rate and virtually trace out Ms. Popovich's supply curve for labor. But, if we pay Ms. Popovich too much, she may actually choose to work less. She has attained enough income to make her comfortable and now wants to spend more time with her family. So if the wage rate rises to W'' , Ms. Popovich will decrease her working time to $24 - H_L$ hours.

^a Arguably, it might be the other way around.

Ms. Popovich's Dilemma of Balancing Work and Leisure

Ms. Popovich gets utility from work (income) and leisure (family and friends). If the wage rate is W , she will work $24 - H_L$ hours. If the wage rate increases to W' , she will increase her work hours to $24 - H'_L$.

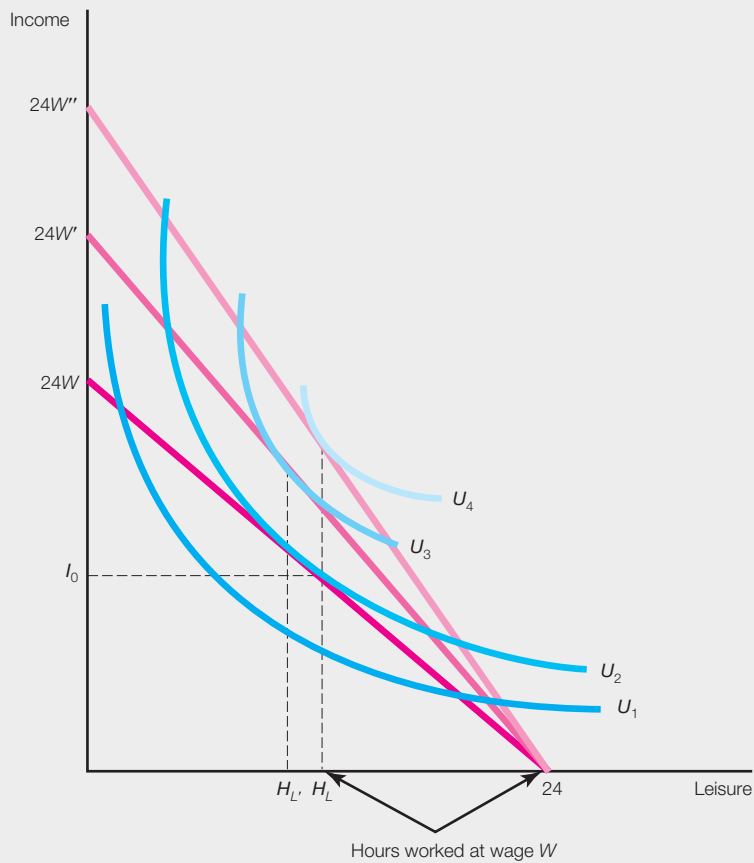
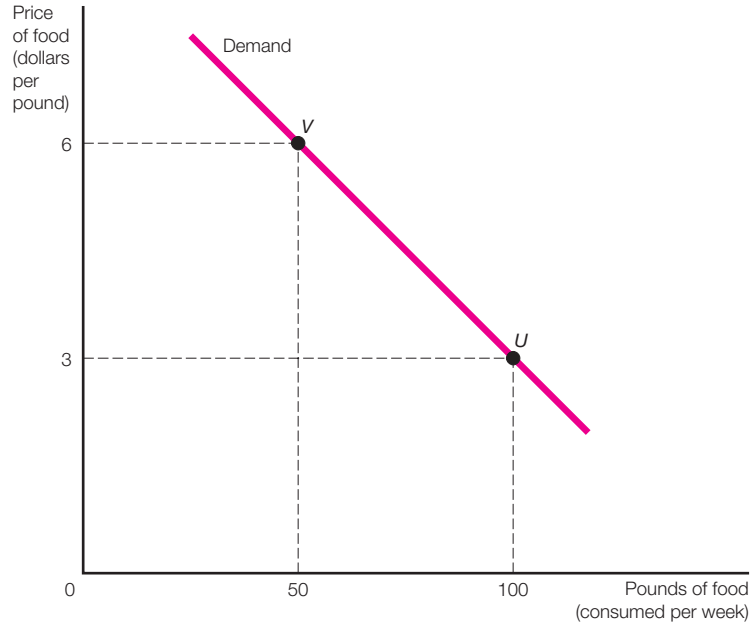


FIGURE 3.14

Ms. Popovich's Individual Demand Curve for Food

Ms. Popovich's individual demand curve for food shows the amount of food she would buy at various prices.

**DERIVING THE MARKET DEMAND CURVE**

We have just shown how to estimate a consumer's demand curve for a product, given the consumer's tastes and income as well as the prices of other products. What if we estimate the individual demand curve for each consumer in the market? How can they help us understand the market?

The answer is easy. Think of the market demand curve as representing the sum of tastes and preferences of individual consumers. It summarizes the demand curves of all individuals in the market. To derive the market demand curve, we estimate the horizontal sum of all the individual demand curves. At each pricing point we estimate the market total by summing the purchases of all individuals at that price.

Table 3.1 shows the demand schedules for food of four families: the Moores, Sarafians, Chases, and Grubers. For simplicity, suppose these four families constitute the entire market for food; then the market demand curve for food is shown in the last column of Table 3.1. Figure 3.15 shows the families' demand curves for food as well as the resulting market demand curve. To illustrate how the market demand curve is derived, suppose the price of food is \$3 per pound; the market quantity demanded is 103 hundreds of pounds per month. This is the sum of the

TABLE 3.1

Individual Demand Curves and Market Demand Curve for Food

Price of Food (Dollars per Pound)	Individual Demand Curves (Hundreds of Pounds per Month)				Market Demand
	Moore	Sarafian	Chase	Gruber	
3.00	51.0	45.0	5.0	2.0	103
3.20	43.0	44.0	4.2	1.8	93
3.40	36.0	43.0	3.4	1.6	84
3.60	30.0	42.0	2.6	1.4	76
3.80	26.0	41.4	2.4	1.2	71
4.00	21.0	41.0	2.0	1.0	65

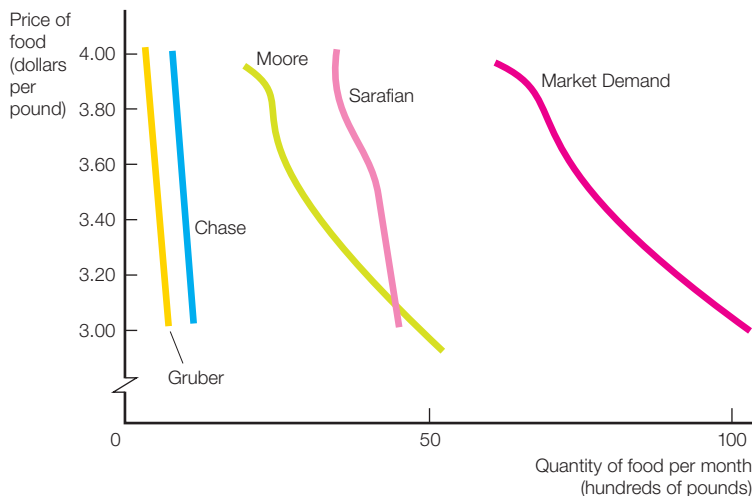
quantities demanded by the four families. (As shown in Table 3.1, this sum equals $51.0 + 45.0 + 5.0 + 2.0$, or 103 hundreds of pounds.)

Figure 3.15 clearly illustrates that within a single product market, demand is not composed of homogeneous buyers. A market is generally composed of buyers with different tastes and preferences. We will see later that managers can strategically exploit this heterogeneity by identifying submarkets and charging each submarket a different price.

FIGURE 3.15

Individual Demand Curves and Market Demand Curve for Food

The market demand curve is the horizontal sum of all the individual demand curves.



We also see why managers like to expand their markets. As more consumers enter the market, the demand curve is pushed out to the right due to horizontal summation. As this occurs (and the supply curve remains constant), the market price increases.

Finally, while demand for a firm's product or service is necessarily a fraction of total market demand, the demand curve facing managers of a firm is generally not a parallel, scaled-down version of market demand. The choices of managers can significantly influence demand for their products. Managers who are better able to influence consumers realize higher performance relative to rivals.

CONSUMER SURPLUS

One key insight managers need to understand is that many consumers value a product at a premium. They are willing to pay a higher price than that of the market, as illustrated in Figure 3.16. This is a simple but powerful (and potentially profitable) fact that managers need to exploit. An individual's demand curve estimates the unit price (P_X) she is willing to pay to purchase a given number of product units (say X). Because the demand curve (usually) is downward-sloping, the curve indicates that the consumer values products $X - 1$, $X - 2$, and so forth at a higher amount than the value of the X th good purchased (say at values P_{X-1} , P_{X-2}). The price at which the consumer values each number of units demanded is called the **consumer's reservation price** for that particular unit of good. It is also known as the willingness to pay (WTP). The reservation price is the highest price the consumer is willing to pay for that unit of product or service. If we try to charge any price above the WTP, the consumer will not purchase from us.

The difference between what an individual is willing to pay and what that individual has to pay (the market price) for a product is called **consumer surplus**. It is the actual price paid subtracted from the reservation price.

When a market is at equilibrium, the marginal individuals to purchase are those whose reservation price just equals the market price. They receive no consumer surplus from the purchase; they value it for the amount of money they paid for it. But all other purchasers have reservation prices exceeding the market price. They all gained a surplus because they paid less for the product than they were willing to pay (they got what many term a "good deal"). If we aggregate all the individuals' consumer surpluses, we estimate the consumer surplus of the market at that given price. Visually, a product's consumer surplus is the area below the demand curve but above the market price (area A in Figure 3.16).

We will say more about consumer surplus later, when we introduce the analogous idea of producer surplus, and the summation of consumer and producer surplus, called the total surplus. Economists use these concepts to describe the efficiency of markets and the social benefits of market transactions. We draw a simpler observation. As long as the demand curve for a product is downward-sloping

Consumer's reservation price

The price at which the consumer values each number of units demanded.

Consumer surplus The difference between what an individual is willing to pay (their reservation price) and what that individual has to pay (the market price) for a product.

STRATEGY SESSION: The Trade-Off between Risk and Return

As investors in developing countries like India and China begin to diversify their savings from simple saving accounts to bonds and stocks, they will show behavior regarding their risk–return profiles. A recent article in *The Times of India* discussed the need for investors to understand their “risk profile.”^a We can operationalize this risk profile using indifference curves.

Assume our investor Devi Bangerjee has \$1 million, which she must allocate between stocks and government bonds. If she invests in government bonds, she will receive a return of 5%, and there is no risk. If she invests \$1 million in common stock, she expects a return of 10% and endures considerable risk. If she invests half in bonds and half in common stocks, she expects a return of 7.5% and there is some risk. Line *RT* in the graph that follows shows her expected return with its corresponding risk for combinations of the two investments. People differ in their risk tolerance; this is part of human nature. We represent hers in the form of indifference curves.

But because we differ in risk tolerance does not mean we are totally idiosyncratic. For most of us,

indifference curves slope upward to the right. Risk is fundamentally different from purchase decisions, for which indifference curves slope downward to the right.

The risk indifference curves slope upward to the right because Devi prefers less risk to more when the expected return is held constant. If there is an increase in risk, she needs a higher expected return to maintain the same level of satisfaction. She must choose some point on line *RT*. The point on *RT* that is tangent with the highest indifference curve is point *S*. Here her expected return is 7.5%. Hence she should purchase \$500,000 of government bonds and \$500,000 of common stock. We will look at such investment decisions in greater detail in Chapter 14.

There are websites around the globe (www.amp.co.nz; www.tools.asiapacific.hsbc.com) with risk profile calculators. Using a short series of questions, these calculators approximate the indifference curve of the investor.

^aD. Ghosh, “Know Your Appetite for Risk-Taking,” *The Times of India*, April 15, 2008.

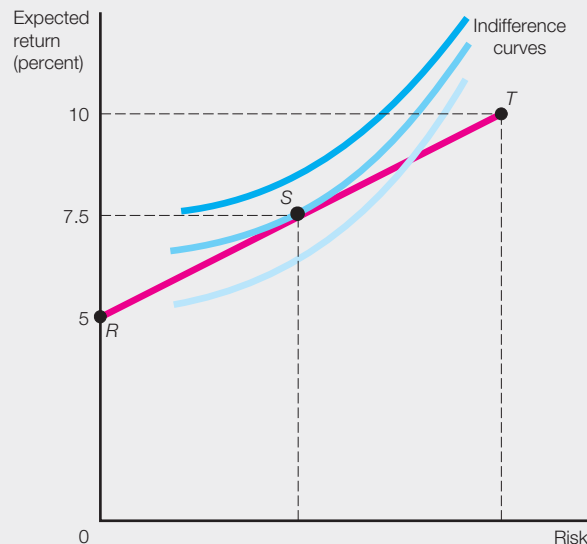
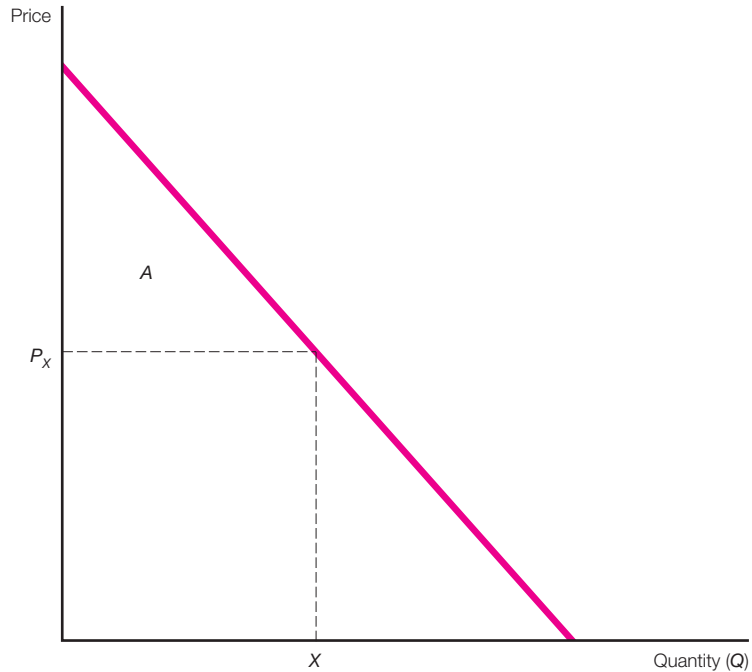


FIGURE 3.16

The Consumer Surplus for a Price of P_x

The consumer surplus for an individual is the area under the demand curve but above the price (P_x) paid (area A). The same definition holds for a market demand curve.



(and most are), a manager will generate more revenue if she charges each consumer his or her reservation price. Managers can also increase revenue by charging higher prices to consumers who value the product more highly. This is called price discrimination. Relative to a simple single-price strategy, this captures consumer surplus for the benefit of the firm. Even though various legal, practical, and economic constraints limit the extent to which managers can charge different prices for the same product, the practice is widespread. Examples include airline tickets, automobiles sold at dealerships where haggling is the norm, and goods and services offered at discounts through coupon systems and other special offers. We explore these strategies in greater detail later.

SUMMARY

1. An indifference curve contains points representing market bundles among which the consumer is indifferent. If the consumer prefers more to less of both commodities, an indifference curve must have a negative slope.

2. Market bundles on higher indifference curves provide more satisfaction than those on lower indifference curves. Utility is a number that indexes the level of satisfaction derived from a particular market bundle. Market bundles with higher utilities are preferred over those with lower utilities.

3. The marginal rate of substitution shows how many units of one good must be given up if the consumer, after getting an extra unit of another good, is to maintain a constant level of satisfaction. To obtain the marginal rate of substitution, multiply the slope of the indifference curve by -1 .

4. The budget line contains all the market bundles the consumer can buy, given his or her money income and the level of each price. Increases in income push the budget line upward and parallel to the old budget line; changes in the price ratio alter the budget line's slope.

5. To attain the highest level of satisfaction compatible with the budget line, the consumer must choose the market bundle on the budget line that is on the highest indifference curve. This market bundle is at a point where the budget line is tangent to an indifference curve (unless there is a corner solution).

6. The consumer who maximizes utility will choose in equilibrium to allocate his or her income so that the marginal rate of substitution of one good for another good equals the ratio of the prices of the two goods (unless there is a corner solution).

7. The theory of consumer behavior is often used to represent the process of rational choice. Frequently a person or organization has a certain amount of money to spend and must decide how much to allocate to a number of different uses. This theory indicates how such decisions should be made.

8. A consumer's demand curve shows how much the consumer would purchase of a good at various prices of the good when other prices, preferences, and the consumer's income are held constant. The theory of consumer behavior can be used to derive the consumer's demand curve, and the market demand curve can be obtained by summing the individual demand curves horizontally.

9. Consumer surplus is the difference between what a consumer is willing to pay for a good and what the consumer pays for the good in the market. Clever managers want to figure out pricing policies to extract consumer surplus from consumers.

PROBLEMS

1. The market for sports performance drinks experienced a big shift in 2012 as sales of low-calorie sports drinks grew by over 25%. Many attributed this shift to greater use by women who wanted a sports drink without many calories.
 - a. If a woman desires two containers of low-calorie sports drink as much as one container of high-calorie sports drink, what do her

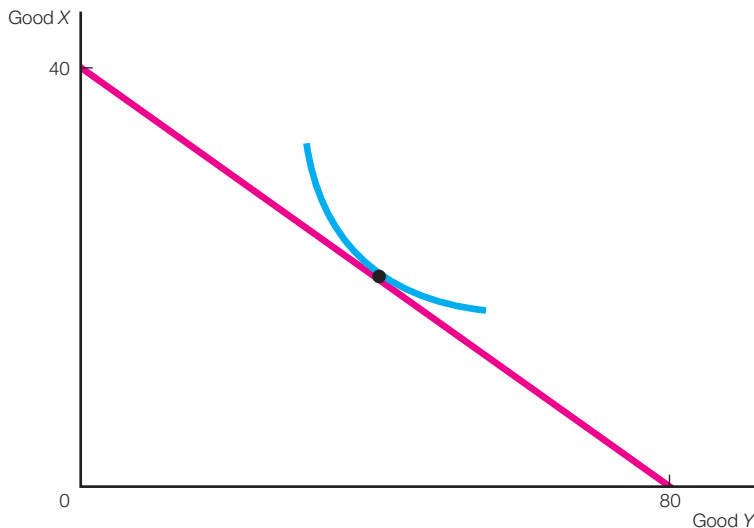
indifference curves (between low- and high-calories sports drinks) look like?

- b. Do they have the typical shape of indifference curves? Why or why not?
2. In recent years fresh bagel sales have been growing at about 30% per year. Once considered an ethnic food to be eaten with cream cheese and lox, bagels now “have become the new donut to bring to the office,” according to Michael Goldstein of Goldstein’s Bagel Bakery in Pasadena, California. But one problem with bagels is that they get stale fast. In the words of Ray Lahvic, editor emeritus of *Bakery Production and Marketing*, “the worst thing in the world is a day-old bagel.” If a market researcher asserts that the slope of the typical consumer’s indifference curves between fresh bagels and day-old bagels is -1 , would you agree with this assertion? Why or why not?
3. On a piece of graph paper, plot the quantity of lamb consumed by Ms. Turner along the vertical axis and the quantity of rice she consumes along the horizontal axis. Draw the indifference curve that includes the following market bundles. Each of these market bundles gives equal satisfaction

Market Bundle	Lamb (Pounds)	Rice (Pounds)
1	2	8
2	3	7
3	4	6
4	5	5
5	6	4
6	7	3
7	8	2
8	9	1

4. In the previous question, what is the marginal rate of substitution of rice for lamb? How does the marginal rate of substitution vary as Ms. Turner consumes more lamb and less rice? Is this realistic?
5. Suppose Richard has an after-tax income of \$500 per week and must spend it all on food or clothing. If food is \$5 per pound and clothing is \$10 per piece, draw his budget line on a piece of graph paper, where the amount of food is measured along the vertical axis and the amount of clothing is measured along the horizontal axis.
6. In the previous problem, what is the budget line if Richard’s weekly income increases to \$600? What is his budget line if his income is \$500, but the price of food increases to \$10 per pound? What is his budget line if his income is \$500, but the price of clothing increases to \$20 per piece? Draw each of these budget lines on the piece of graph paper used in the previous problem.

7. Maria has budgeted a total of \$9 to spend on two goods: chips and salsa. She likes to consume a unit of chips in combination with a unit of salsa. Any unit of chips that she cannot consume in combination with a unit of salsa is useless. Similarly, any unit of salsa that she cannot consume in combination with a unit of chips is useless. If the price of a unit of chips is \$0.50 and the price of a unit of salsa is \$0.10, how many units of each good does she purchase?
8. In the following diagram, we show one of Jane's indifference curves and her budget line.
 - a. If the price of good X is \$100, what is her income?
 - b. What is the equation for her budget line?
 - c. What is the slope of her budget line?
 - d. What is the price of good Y?
 - e. What is Jane's marginal rate of substitution in equilibrium?



9. Sarah has \$300 to allocate between opera tickets and movie tickets. The price of each opera ticket is \$60, and the price of each movie ticket is \$6. Her marginal rate of substitution of opera tickets for movie tickets equals 5, regardless of what market bundle she chooses. How many opera tickets does she purchase?
10. Suppose Milton has \$50 to be divided between corn and beans and that the price of beans is \$0.50 per pound. What will be the relationship between the price of corn and the amount of corn he will buy if $U = \log Q_c + 4 \log Q_b$, where U is his utility, Q_c is the quantity of corn he consumes (in pounds), and Q_b is the quantity of beans he consumes (in pounds)?
11. The state of New York receives \$3 billion (from federal sources and a state petroleum tax) to be spent on highways and/or mass transit (subways, buses,

and urban rail lines), both of which could be used to meet the transportation needs of the state's population.

- a. If each mile of mass transit costs \$20 million, what is the maximum number of miles of mass transit that these funds would have enabled the state to construct?
- b. If each mile of highway costs \$10 million, what is the maximum number of miles of highways that these funds would have enabled the state to construct?
- c. If the number of miles of mass transit constructed is put on the vertical axis of a graph and the number of miles of highways constructed is put on the horizontal axis, can a budget line (showing the maximum number of miles of mass transit that can be constructed, given each number of miles of highways constructed) be drawn for the state? If so, what is the slope of this budget line? (Assume that the \$3 billion is the only source of funds for mass transit or highway construction.)
- d. If the public and the state government agree that every extra mile of mass transit adds three times as much to the state's transportation capability as an extra mile of highways, how much of the \$3 billion should be spent on mass transit if the objective is to maximize transportation capability?

ESTIMATING DEMAND FUNCTIONS

Richard Williams, marketing director at Verizon Wireless, was considering a shift in marketing dollars from traditional marketing channels—newspaper, TV, and the web—to mobile ads and social media.

All firms, not just Verizon, must constantly reevaluate and adjust their marketing strategies to formulate an effective marketing strategy. As stressed in previous chapters, a manager must have a good working knowledge of the demand for his or her firm's products.

The previous two chapters were concerned with the theory of demand; now we investigate techniques to estimate a product's demand function. Consumer surveys and market experiments are useful in providing such information, but the technique most frequently used to estimate demand is regression analysis.

In Chapter 2, we showed how Amtrak managers estimated demand with regression analysis (see page 54). Since regression analysis is used repeatedly in subsequent chapters to estimate production functions and cost functions and for forecasting, we devote considerable attention to this basic technique in this chapter.

THE IDENTIFICATION PROBLEM

While it is important that managers have reasonably accurate estimates of the demand functions for their own (and other) products, this does not mean that it is always easy to obtain such estimates. One problem that may arise in estimating

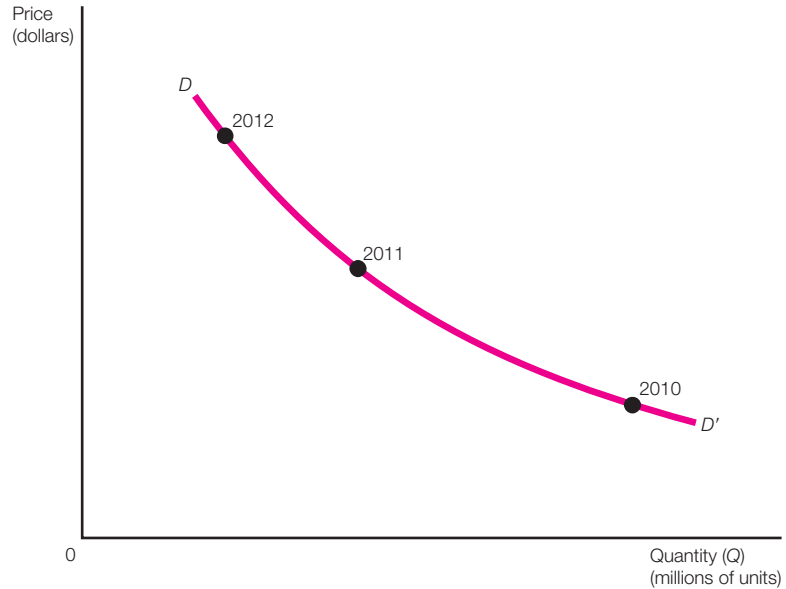
LEARNING OBJECTIVES

- The Identification Problem
- Consumer Interviews
- Market Experiments
- Regression Analysis
- Simple Regression Model
- Sample Regression Line
- Method of Least Squares
- Coefficient of Determination
- Multiple Regression
- Software Packages and Computer Printouts
- Interpreting the Output of Statistical Software
- Multicollinearity
- Serial Correlation
- Further Analysis of the Residuals
- Summary
- Problems
- Appendix: The Coefficient of Determination and the Concept of Explained Variation

FIGURE 4.1

Price Plotted against Quantity, 2010–2012

The curve DD' is unlikely to be a good estimate of the demand curve.



demand should be recognized at the outset. Given the task of estimating the demand curve for a particular product, managers may be inclined to plot the quantity demanded of the product in 2012 versus its 2012 price, the quantity demanded in 2010 versus its 2010 price, and so forth. If the resulting plot of points for 2010 to 2012 were as shown in Figure 4.1, a manager might be tempted to conclude that the demand curve is DD' .

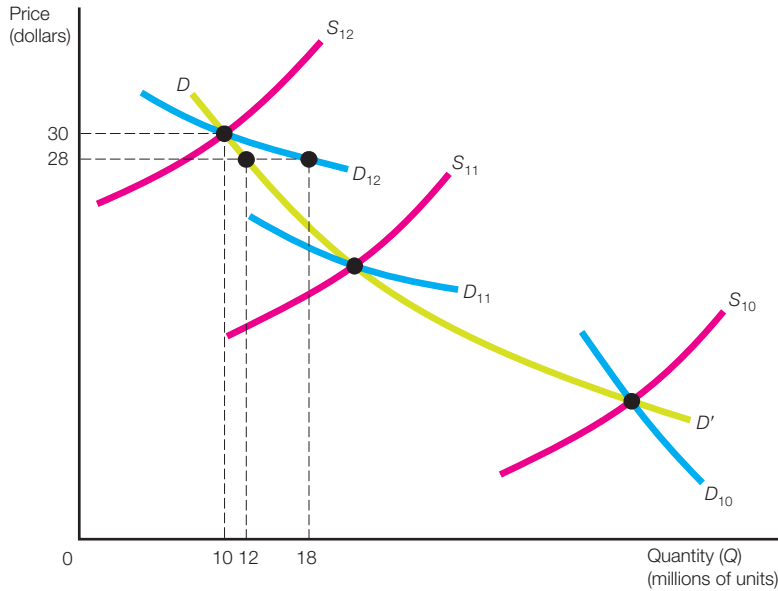
Unfortunately, things are not so simple. Price, as we saw in, Chapter 1 is determined by both the demand and supply curves for this product if the market is competitive. Specifically, the equilibrium value of price is at the level where the demand and supply curves intersect. The important point to note is that the demand and supply curves for this product may be different each year. So, as shown in Figure 4.2, the supply curve may have shifted (from S_{10} in 2010 to S_{11} in 2011 to S_{12} in 2012), and the demand curve may have shifted (from D_{10} in 2010 to D_{11} in 2011 to D_{12} in 2012). As indicated in Figure 4.2, DD' is not even close to being a good approximation to the demand curve for this product in any of these three years.

In the situation in Figure 4.2, if you were to conclude that DD' was the demand curve, you would underestimate (in absolute value) the price elasticity of demand for this product in 2012 and 2011 and overestimate it (in absolute value) in 2010. In 2012, you would think that, if price were lowered from \$30 to \$28, the

FIGURE 4.2

Estimated Demand Curve Contrasted with Actual Demand Curves

The estimated demand curve DD' is not at all similar to the actual demand curves.



quantity demanded would increase from 10 to 12 million units per year. In fact, as shown in Figure 4.2, such a price reduction would result in an increase of the quantity demanded to 18, not 12, million units per year. This is a mammoth error in anyone's book.

The point is that, because a manager is not holding constant a variety of non-price variables like consumer tastes, incomes, the prices of other goods, and advertising, she cannot be sure that the demand curve was fixed during the period when the measurements were made. If the demand curve was fixed and only the supply curve changed during the period, she could be confident that the plot of points in Figure 4.1 represents the demand curve. As shown in Figure 4.3, the shifts in the supply curve trace out various points on the demand curve we want to measure.

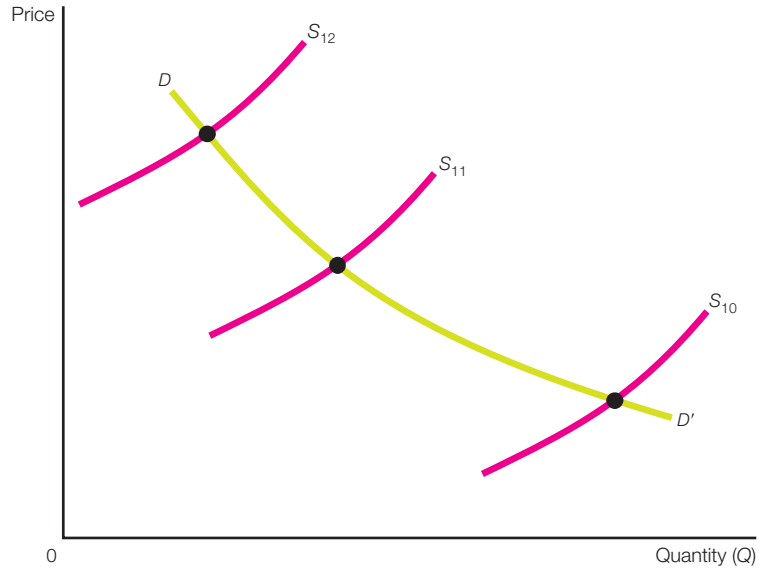
How can managers estimate a demand curve if it has not remained fixed in the past? There are many ways, some simple, some very complex. Econometric techniques recognize that price and quantity are related by both the supply curve and the demand curve and both these curves shift in response to nonprice variables. Some basic econometric techniques, such as regression analysis, are presented later in this chapter; others are too complex to be taken up here.¹ Consumer interviews and market experiments are also widely used, as indicated in the next three sections.

1. See J. Johnston and J. DiNardo, *Econometric Methods* (4th ed.; New York: McGraw-Hill, 1997); J. Kmenta, *Elements of Econometrics* (3rd ed.; New York: Macmillan Co., 1997); or E. Berndt, *The Practice of Econometrics* (2nd ed.; Reading, MA: Addison-Wesley, 1996).

FIGURE 4.3

Fixed Demand Curve and Shifting Supply Curve

In this special case, DD' does represent the actual demand curve.

**CONSUMER INTERVIEWS**

To obtain information concerning the demand function for a particular product, managers frequently interview consumers and administer questionnaires concerning their buying habits, motives, and intentions. Managers may also run focus groups in an attempt to discern consumers' tastes. For example, a manager might ask a random sample of consumers how much more gasoline they would purchase if its price were reduced by 5%. Or, a market researcher might ask a sample of consumers whether they liked a new type of tablet better than a leading existing brand, and if so, how much more they would be willing to pay for it (than for the existing brand).

Unfortunately, consumer surveys of this sort have many well-known limitations. The direct approach of simply asking people how much they would buy of a particular commodity at particular prices often does not seem to work very well. Frequently, the answers provided by consumers to such a hypothetical question are not very accurate. However, more subtle approaches can be useful. For example, in one study interviews indicated that most buyers of a particular baby food selected it on their doctor's recommendation and that most of them knew very little about prices of substitutes. This information, together with other data, suggested that the price elasticity of demand was quite low in absolute value.

Despite the limitations of consumer interviews and questionnaires, many managers believe that such surveys can reveal a great deal about consumer preferences. For example, one survey conducted at the Wharton School by a leading skin care company found that coeds did not want to purchase a skin cream when they were told that it eliminated wrinkles, but would purchase it when told it made the skin glow.

MARKET EXPERIMENTS

Another method of estimating demand for a particular commodity is to carry out direct market experiments. The idea is to vary the price of the product while attempting to keep other market conditions fairly stable (or to take changes in other market conditions into account). For example, managers of a cell phone company conducted an experiment some years ago to determine the price elasticity of demand for its product. They raised the price of their cell phones by 15% in four cities and found that demand was quite inelastic. Attempts were made to estimate the cross elasticity of demand with other brands as well.

Controlled laboratory experiments can sometimes be carried out. Consumers are given money and told to shop in a simulated store. The experimenter can vary the prices, packaging, and location of particular products, and see the effects on the consumers' purchasing decisions. While this technique is useful, it suffers from the fact that consumers participating in such an experiment know that their actions are being monitored. For that reason, their behavior may depart from what it normally would be.

Before carrying out a market experiment managers need to weigh the costs against the benefits. Direct experimentation can be expensive or risky because customers may be lost and profits cut by the experiment. For example, if the price of a product is raised as part of an experiment, potential buyers may be driven away. Also, since they are seldom really controlled experiments and since they are often of relatively brief duration and the number of observations is small, experiments often cannot produce all the information that is needed.

REGRESSION ANALYSIS

Although consumer interviews and market experiments are important sources of information about consumer preferences, they are of limited value in regression analysis. Suppose a firm's demand function is

$$Y = A + B_1X + B_2P + B_3I + B_4P_r \quad (4.1)$$

where Y is the quantity demanded of the firm's product, X is the marketing expense (such as advertising), P is the price, I is the disposable income of consumers, and P_r is the price of rival brands. What managers need are estimates of the values of A ,

STRATEGY SESSION: Marketing Plans at the Stafford Company

The Stafford Company developed a new type of electric drive. When the design engineering for this machine was finished, Stafford's managers began to make long-range plans concerning marketing this product. By means of field surveys and the analysis of published information, the firm's market research personnel estimated that about 10,000 electric drives of this general sort would be sold per year. The share of the total market that Stafford's new product would capture depended on its price. According to the firm's market research department, the relationship between price and market share was as follows:

Price	Market Share
\$ 800	11.0
900	10.2
1,000	9.2
1,100	8.4
1,200	7.5
1,300	6.6
1,400	5.6

Stafford's managers wanted advice in setting the price for their new drive, and to help determine the optimal price, they wanted a simple equation expressing the annual quantity demanded of the new product as a function of its price. They also wanted whatever information could readily be provided concerning the reliability of this equation. In particular, they were interested in whether they could safely use this equation to estimate the quantity demanded if price were set at \$1,500 or \$1,600.

Prepare a brief report supplying the information requested. (Note that the figures on market share in the table are expressed in percentage points. Thus, if the price of Stafford's new product is set at \$800, it will capture 11.0 % of the market for electric drives of this general sort, according to the market research department.)

Source: This section is based on an actual case, although the numbers and situation are disguised somewhat.

B_1 , B_2 , B_3 , and B_4 . Regression analysis enables us to obtain them from historical data concerning Y , X , P , I , and P_r .

In the rest of this chapter, we describe the nature and application of regression analysis, a statistical technique used to estimate many types of economic relationships, not just demand functions. We begin with the basic case in which marketing expense is the only factor to significantly affect demand. We then turn to the more complicated (and realistic) case in which the quantity demanded is affected by more than one factor, as in equation (4.1).

Regression analysis describes the way in which one variable is related to another. (While the technique can handle more than two variables, only two are considered here.) Regression analysis derives an equation used to estimate the unknown value of one variable on the basis of the known value of another. For example, suppose

TABLE 4.1

**Marketing Expense and Sales, Miller Pharmaceutical Company,
Sample of Nine Years**

Marketing Expense (Millions of Dollars)	Sales (Millions of Units)
1	4
2	6
4	8
8	14
6	12
5	10
8	16
9	16
7	12

the Miller Pharmaceutical Company is scheduled to spend \$4 million next year on marketing (for promotion, advertising, etc.) and managers need to estimate its next-year's sales, on the basis of the historical data in Table 4.1. In this case, although the firm's marketing expense next year is known, its next year's sales are unknown. Regression analysis describes the way in which the sales are historically impacted by marketing.

SIMPLE REGRESSION MODEL

As you recall from Chapter 1, a **model** is a simplified or idealized representation of reality. In this section, we describe the model—that is, the set of simplifying assumptions—on which regression analysis is based. We begin by visualizing a population of all relevant pairs of observations of the independent and dependent variables. For instance, in the case of the Miller Pharmaceutical Company, we visualize a population of pairs of observations concerning sales and marketing expense. This population includes all the levels of sales corresponding to all the levels of marketing in the history of the firm.

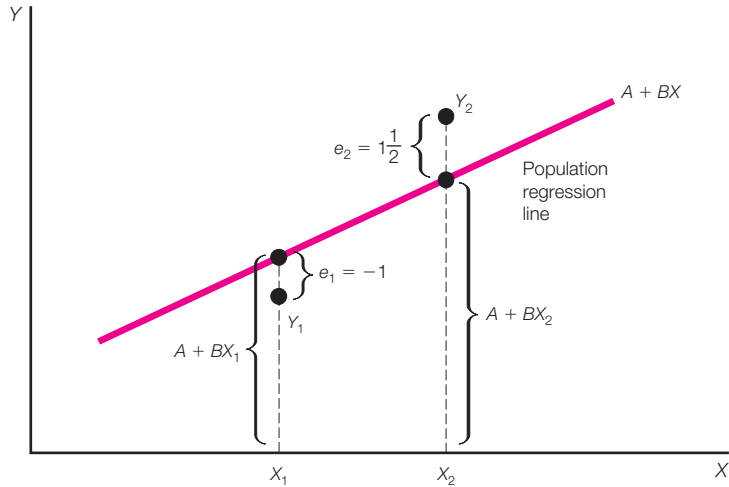
The mean of a variable equals the sum of its values divided by their number. Therefore, the mean of a variable that assumes four values, 3, 2, 1, and 0, is $(3 + 2 + 1 + 0)/4$, or 1.5. Regression analysis assumes that *the mean value of Y, given the value of X, is a linear function of X*. In other words, the mean value of the dependent variable is assumed to be a linear function of the independent variable,

Model A simplified or idealized representation of reality.

FIGURE 4.4

Regression Model

The mean value of Y , given the value of X , falls on the population regression line.

**Population (true) regression line**

The line resulting from regressing the dependent variable on the independent variable where the entire population of values of the variables is used.

Error term A random amount that is added or subtracted from the population regression line.

2. The values of e_1 and e_2 are independent if the probability distribution of e_1 does not depend on the value of e_2 and the probability distribution of e_2 does not depend on the value of e_1 . Regression analysis also assumes that the variability of the values of e_i is the same, regardless of the value of X . Many of the tests described subsequently assume too that the values of e_i are normally distributed. For a description of the normal distribution, see Appendix E.

the equation of this being $A + BX$, as shown in Figure 4.4. This straight line is called the **population regression line** or the **true regression line**.

Put differently, regression analysis assumes that

$$Y_i = A + BX_i + e_i \quad (4.2)$$

where Y_i is the i th observed value of the dependent variable and X_i is the i th observed value of the independent variable. Essentially, e_i is an **error term**, that is, a random amount that is added to $A + BX_i$ (or subtracted from it if e_i is negative). Because of the presence of this error term, the observed values of Y_i fall around the population regression line, not on it. Hence, as shown in Figure 4.4, if e_1 (the value of the error term for the first observation) is -1 , Y_1 lies 1 below the population regression line. And if e_2 (the value of the error term for the second observation) is $+1.50$, Y_2 lies 1.50 above the population regression line. Regression analysis assumes that the values of e_i are independent and their mean value equals zero.²

Although the assumptions underlying regression analysis are unlikely to be met completely, they often closely match reality. So regression analysis is a powerful technique. Nonetheless, the validity of the estimates depends on the congruency of assumptions with reality.

SAMPLE REGRESSION LINE

The purpose of a regression analysis is to obtain the mathematical equation for a line that describes the average relationship between the dependent and independent variables. This line is calculated from the sample observations and is called the **sample** or **estimated regression line**. It should not be confused with the population regression line discussed in the previous section. Whereas the population regression line is based on the entire population, the sample regression line is based on only the sample.

The general expression for the sample regression line is

$$\hat{Y} = a + bX$$

where \hat{Y} is the value of the dependent variable predicted by the regression line, and a and b are estimators of A and B , respectively. (An estimator is a function of the sample observations used to estimate an unknown parameter. For example, the sample mean is an estimator often used to estimate the population mean.) Since this equation implies that $\hat{Y} = a$ when $X = 0$, it follows that a is the value of \hat{Y} at which the line intersects the Y axis. Therefore, a is often called the **Y intercept of the regression line**. And b , which clearly is the slope of the line, measures the change in the predicted value of Y associated with a one-unit increase in X .

Figure 4.5 shows the estimated regression line for the data concerning sales and marketing expense of the Miller Pharmaceutical Company. The equation for this regression line is

$$\hat{Y} = 2.536 + 1.504X$$

where \hat{Y} is sales in millions of units and X is marketing expense in millions of dollars. What is 2.536? It is the value of a , the estimator of A . What is 1.504? It is the value of b , the estimator of B . For the moment, we are not interested in how this equation was determined; what we want to consider is how it should be interpreted.

At the outset, note the difference between Y and \hat{Y} . Whereas Y denotes an *observed* value of sales, \hat{Y} denotes the *computed* or *estimated* value of sales, based on the regression line. For example, the first row of Table 4.1 shows that, in the first year, the actual value of sales was 4 million units when marketing expense was \$1 million. Therefore, $Y = 4.0$ millions of units when $X = 1$. In contrast, the regression line indicates that $\hat{Y} = 2.536 + 1.504(1)$, or 4.039 millions of units when $X = 1$. In other words, while the regression line predicts that sales will equal 4.039 millions of units when marketing expense is \$1 million, the actual sales figure under these circumstances (in the first year) was 4 million units.

It is essential to be able to identify and interpret the Y intercept and slope of a regression line. What is the Y intercept of the regression line in the case of the

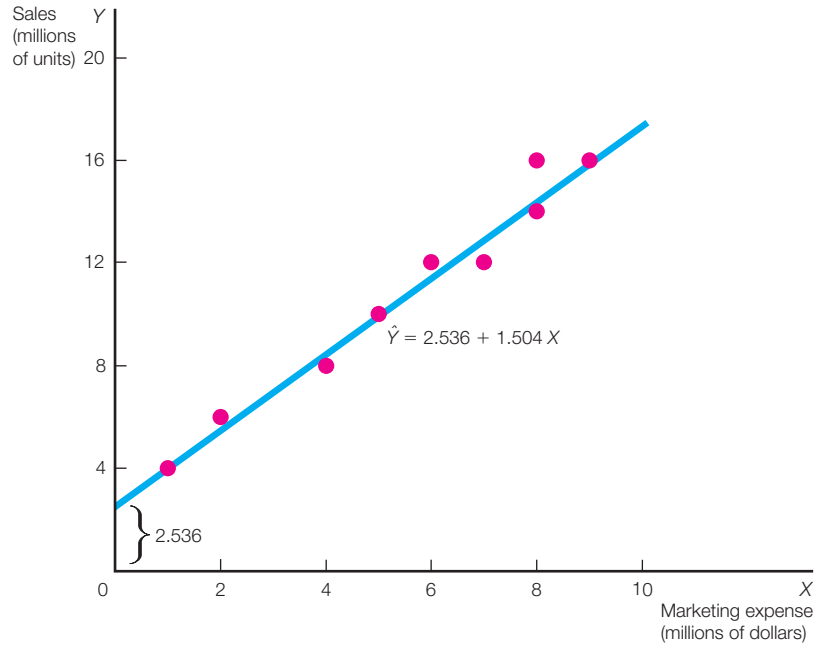
Sample (estimated) regression line The line resulting from regressing the dependent variable on the independent variable where only a sample of the variables is used.

Intercept of the regression line The value of the dependent variable when the independent variables have a value of zero.

FIGURE 4.5

Sample Regression Line

This line is an estimate of the population regression line.



Miller Pharmaceutical Company? It is 2.536 millions of units. This means that, if managers allocate \$0 to marketing, the estimated sales would be 2.536 millions of units. (As shown in Figure 4.5, 2.536 millions of units is the value of the dependent variable at which the regression line intersects the vertical axis.) What is the slope of the regression line in this case? It is 1.504. This means that the estimated sales go up by 1.504 millions of units when marketing expense increases by \$1 million.

3. As pointed out in Chapter 1, Σ is the mathematical summation sign. What does ΣX_i mean? It means that the numbers to the right of the summation sign (that is, the values of X_i) should be summed from the lower limit on i (which is given below the Σ sign) to the upper limit on i (which is given above the Σ sign)

$$\sum_{i=1}^n X_i$$

means the same thing as $X_1 + X_2 + \cdots + X_n$.

METHOD OF LEAST SQUARES

The method used to determine the values of a and b is the so-called method of least squares. Since the deviation of the i th observed value of Y from the regression line equals $\hat{Y}_i - Y_i$, the sum of these squared deviations equals

$$\sum_{i=1}^n (Y_i - \hat{Y}_i)^2 = \sum_{i=1}^n (Y_i - A - BX_i)^2 \quad (4.3)$$

where n is the sample size.³ Using the minimization technique presented in Chapter 18, we can find the values of a and b that minimize the expression in equa-

tion (4.3) by differentiating this expression with respect to a and b and setting these partial derivatives equal to zero

$$\frac{\partial \sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\partial a} = -2 \sum_{i=1}^n (Y_i - a - bX_i) = 0 \quad (4.4)$$

$$\frac{\partial \sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\partial b} = -2 \sum_{i=1}^n X_i (Y_i - a - bX_i) = 0 \quad (4.5)$$

Solving equations (4.4) and (4.5) simultaneously and letting \bar{X} equal the mean value of X in the sample and \bar{Y} equal the mean value of Y , we find that

$$b = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2} \quad (4.6)$$

$$a = \bar{Y} - b\bar{X} \quad (4.7)$$

The value of b in equation (4.6) is often called the **estimated regression coefficient**.

From a computational point of view, it frequently is easier to use a somewhat different formula for b than the one given in equation (4.6). This alternative formula, which yields the same answer as equation (4.6), is

$$b = \frac{n \sum_{i=1}^n X_i Y_i - \left(\sum_{i=1}^n X_i \right) \left(\sum_{i=1}^n Y_i \right)}{n \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2}$$

In the case of the Miller Pharmaceutical Company, Table 4.2 shows the calculation of $\sum X_i Y_i$, $\sum X_i^2$, $\sum X_i$, and $\sum Y_i$. Based on these calculations,

$$b = \frac{9(638) - (50)(98)}{9(340) - 50^2} = 1.504$$

Therefore, the value of b , the least-squares estimator of B , is 1.504, which is the result given in the previous section. In other words, an increase in marketing expense of \$1 million results in an increase in estimated sales of about 1.504 millions of units.

Having calculated b , we can readily determine the value of a , the least-squares estimator of A . According to equation (4.7),

$$a = \bar{Y} - b\bar{X}$$

Estimated regression coefficient

The slope of the sample regression line. It, in turn, is an estimate of the slope of the population regression line.

TABLE 4.2

Computation of $\sum X_i$, $\sum Y_i$, $\sum X_i^2$, $\sum Y_i^2$, and $\sum X_i Y_i$

	X_i	Y_i	X_i^2	Y_i^2	$X_i Y_i$
	1	4	1	16	4
	2	6	4	36	12
	4	8	16	64	32
	8	14	64	196	112
	6	12	36	144	72
	5	10	25	100	50
	8	16	64	256	128
	9	16	81	256	144
	7	12	49	144	84
Total	50	98	340	1,212	638

$$\bar{X} = 50/9 = 5.556$$

$$\bar{Y} = 98/9 = 10.889$$

where \bar{Y} is the mean of the values of Y , and \bar{X} is the mean of the values of X . Since, as shown in Table 4.2, $\bar{Y} = 10.889$ and $\bar{X} = 5.556$, it follows that

$$\begin{aligned} a &= 10.889 - 1.504(5.556) \\ &= 2.536 \end{aligned}$$

Therefore, the least-squares estimate of A is 2.536 millions of units, which is the result given in the previous section.

Having obtained a and b , it is a simple matter to specify the average relationship in the sample between sales and marketing expense for the Miller Pharmaceutical Company. This relationship is

$$\hat{Y} = 2.536 + 1.504X \quad (4.8)$$

where \hat{Y} is measured in millions of units and X is measured in millions of dollars. As we know, this line is often called the *sample regression line*, or the *regression of Y on X* . It is the line presented in the previous section and plotted in Figure 4.5. Now, we see how this line is derived.

To illustrate how a regression is used, suppose that the managers of the firm want to predict the firm's sales if they decide to devote \$4 million to marketing expense. Using equation (4.8), they would predict that sales would be

$$2.536 + 1.504(4) = 8.55. \quad (4.9)$$

Since sales are measured in millions of units, this means that sales are expected to be 8.55 million units.

COEFFICIENT OF DETERMINATION

Once the regression line is calculated, managers want to know how well this line fits the data. There can be huge differences in how well a regression line fits a set of data, as shown in Figure 4.6. Clearly, the regression line in panel F of Figure 4.6 provides a much better fit than the regression line in panel B of the same figure. How can managers measure how well a regression line fits the data?

The most commonly used measure of the goodness of fit of a regression line is the coefficient of determination. For present purposes, it is not necessary to know the formula for the coefficient of determination, because it is seldom calculated by hand. It is a particular item, often designated by R^2 , or R-sq on a computer print-out, as we shall see in the section after next.

The value of the **coefficient of determination** varies between 0 and 1. *The closer it is to 1, the better the fit; the closer it is to 0, the poorer the fit.* In the case of the Miller Pharmaceutical Company, the coefficient of determination between sales and marketing expense is 0.97, which indicates a very good fit. To get a feel for what a particular value of the coefficient of determination means, look at the six panels of Figure 4.6. Panel A shows that, if the coefficient of determination is 0, there is no relationship at all between the independent and dependent variables. Panel B shows that, if the coefficient of determination is 0.2, the regression line fits the data rather poorly. Panel C shows that, if it is 0.4, the regression line fits better but not very well. Panel D shows that, if it is 0.6, the fit is reasonably good. Panel E shows that, if it is 0.8, the fit is good. Finally, panel F shows that, if it is 1.0, the fit is perfect.⁴ (A fuller discussion of the coefficient of determination is provided in the appendix to this chapter.)

MULTIPLE REGRESSION

Previously, we discussed regression techniques in the case in which there is only one independent variable. In practical applications of regression techniques, it frequently is necessary and desirable to include two or more independent variables. Now, we extend the treatment of regression to the case in which there is more than one independent variable.

Whereas a **simple regression** includes only one independent variable, a **multiple regression** includes two or more independent variables. Multiple regressions ordinarily are carried out with the aid of statistical software packages. So, there is no reason for managers to learn how to do them by hand. The first step in multiple regression analysis is to identify the independent variables and specify the mathematical form of the equation relating the mean value of the dependent variable to these independent variables.

Coefficient of determination A measure of how well the sample regression line fits the data. Its square root, the correlation coefficient, performs the same function.

Simple regression A dependent variable and one independent variable.

Multiple regression A dependent variable and more than one independent variable.

4. If one is doing the calculations by hand, a convenient formula for the coefficient of determination is

$$r^2 = \frac{\left[n \sum_{i=1}^n X_i Y_i - \left(\sum_{i=1}^n X_i \right) \left(\sum_{i=1}^n Y_i \right) \right]^2}{\left[n \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2 \right] \left[n \sum_{i=1}^n Y_i^2 - \left(\sum_{i=1}^n Y_i \right)^2 \right]}$$

Table 4.2 contains the quantities to be inserted in this formula.

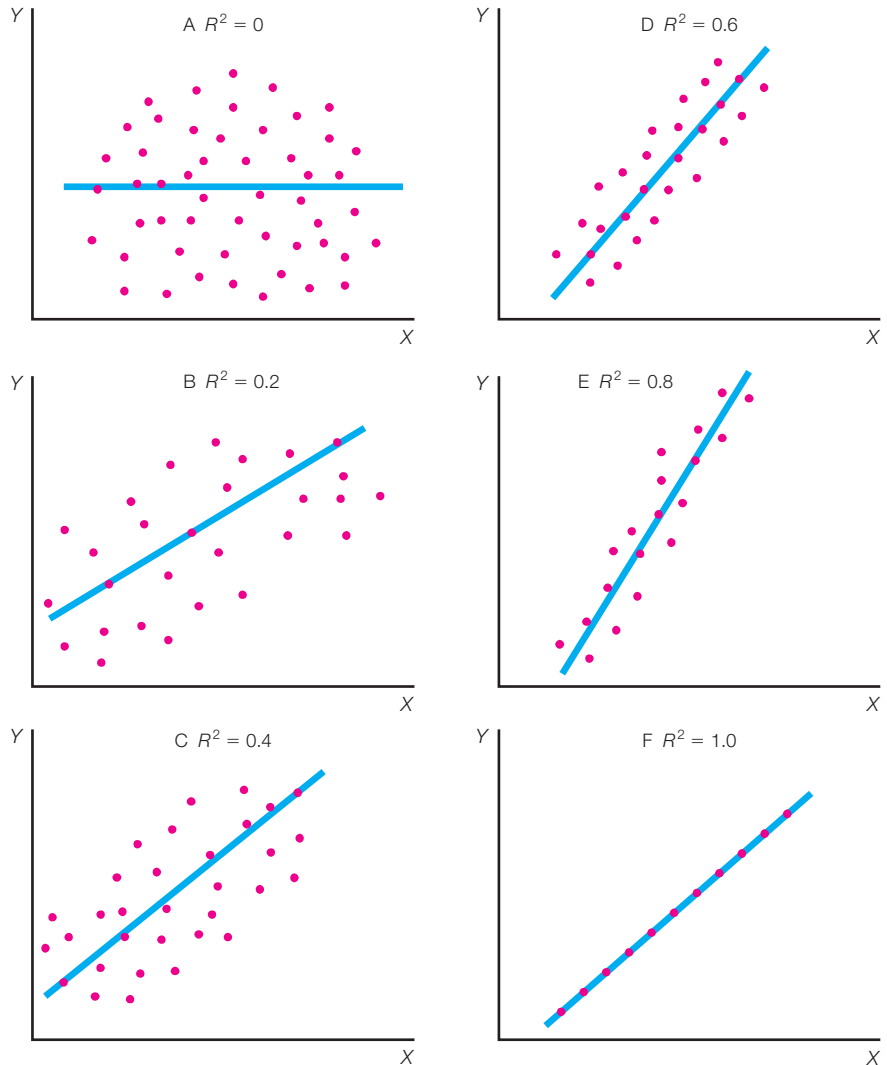
Note too that the square root of r^2 , called the **correlation coefficient**, is also used to measure how well a simple regression equation fits the data. (The sign of the square root is the same as that of b .)

As pointed out in the note to Figure 4.6, computer printouts generally refer to the coefficient of determination as R^2 , although statisticians often call it r^2 when there is only one independent variable.

FIGURE 4.6

Six Regression Lines: Coefficient of Determination Equals 0, 0.2, 0.4, 0.6, 0.8, and 1.0

When there is only one independent variable, the coefficient of determination is often designated by r^2 , rather than R^2 , but computer printouts generally use R^2 , regardless of the number of independent variables. We use R^2 here, even though there is only one independent variable. See footnote 4.



STRATEGY SESSION: How the Japanese Motorcycle Makers Used the Coefficient of Determination

In late 1982, Harley-Davidson asked the International Trade Commission (ITC), a federal agency that investigates possible injuries to U.S. firms and workers from imports, for relief from Japanese imports of heavyweight motorcycles. According to Harley-Davidson, the Japanese were selling their motorcycles at prices too low for it to meet. On the basis of Section 201 of the 1974 Trade Act, the ITC can impose tariffs or quotas on imported goods to provide “additional time to permit a seriously injured domestic industry to become competitive.” But to receive such tariff or quota relief, the industry must demonstrate that the injuries it suffers are due to increased imports, not some other cause such as bad management or a recession.

Harley-Davidson’s petition to the ITC was contested by the major Japanese motorcycle makers: Honda, Kawasaki, Suzuki, and Yamaha. One of their arguments was that general economic conditions, not Japanese imports, were the principal cause of Harley-Davidson’s declining share of the market. In other words, they attributed Harley-Davidson’s problems to the recession of the early 1980s. They pointed out that heavyweight motorcycles, which cost about \$7,000, were a “big-ticket luxury consumer product” and that their sales would be expected to fall in a recession.

To back up this argument, John Reilly of ICF, Inc., the Japanese firms’ chief economic consultant, calculated a regression, where Harley-Davidson’s sales were the dependent variable and the level of blue-collar employment (a measure of general economic

conditions) was the independent variable. He showed that the coefficient of determination was about 0.73. Then, he calculated a regression where Harley-Davidson’s sales were the dependent variable, and the level of sales of Japanese motorcycles was the independent variable. He showed that the coefficient of determination was only about 0.22. From this comparison of the two coefficients of determination, he concluded that Harley-Davidson’s sales were much more closely related to general economic conditions than to the level of sales of Japanese motorcycles.

Of course, this analysis tells us nothing about the effects of the price of Japanese motorcycles on Harley-Davidson’s sales and profits. From many points of view, what was needed was an estimate of the market demand function for Harley-Davidson’s motorcycles. Such an analysis would have related Harley-Davidson’s sales to the price of Harley-Davidson’s motorcycles, the price of Japanese motorcycles, the level of disposable income, and other variables discussed in Chapter 2. In any event, despite the evidence cited, the Japanese motorcycle manufacturers did not prevail. On the contrary, the ITC supported Harley-Davidson’s petition, and on April 1, 1983, President Ronald Reagan imposed a substantial tariff (almost 50%) on imported (large) motorcycles.^a

^aFor further discussion, see J. Gomez-Ibanez and J. Kalt, *Cases in Microeconomics* (Englewood Cliffs, NJ: Prentice-Hall, 1990); P.C. Reid, *Well Made in America: Lessons from Harley-Davidson on Being the Best* (New York: McGraw-Hill, 1989); and *The New York Times*, July 20, 1997.

In the case of the Miller Pharmaceutical Company, suppose that the firm's managers feel that its sales depend on its price, as well as on its marketing expense. More specifically, they assume that

$$Y_i = A + B_1X_i + B_2P_i + e_i \quad (4.10)$$

where X_i is the marketing expense (in millions of dollars) during the i th year and P_i is the price (in dollars) of the product during the i th year (measured as a deviation from \$10, the current price). Of course, B_2 is assumed to be negative. This is a different model from that in equation (4.2). Here, we assume that Y_i (the firm's sales in the i th year) depends on two independent variables, not one. Of course, there is no reason why more independent variables cannot be added, so long as data are available concerning their values and there is good reason to expect them to affect Y_i . But, to keep matters simple, we assume that managers believe that only marketing expense and price should be included as independent variables.⁵

The object of multiple regression analysis is to estimate the unknown constants A , B_1 , and B_2 in equation (4.10). Just as in the case of simple regression, these constants are estimated by finding the value of each that minimizes the sum of the squared deviations of the observed values of the dependent variable from the values of the dependent variable predicted by the regression equation. Suppose that a is an estimator of A , b_1 is an estimator of B_1 , and b_2 is an estimator of B_2 . Then, the value of the dependent variable Y_i predicted by the estimated regression equation is

$$\hat{Y}_i = a + b_1X_i + b_2P_i$$

and the deviation of this predicted value from the actual value of the dependent variable is

$$Y_i - \hat{Y}_i = Y_i - a - b_1X_i - b_2P_i$$

If these deviations are squared and summed, the result is

$$\sum_{i=1}^n (Y_i - \hat{Y}_i)^2 = \sum_{i=1}^n (Y_i - a - b_1X_i - b_2P_i)^2 \quad (4.11)$$

where n is the number of observations in the sample. As pointed out earlier, we choose the values of a , b_1 , and b_2 that minimize the expression in equation (4.11). These estimates are called *least-squares estimates*, as in the case of simple regression.

Software programs are available to calculate these least-squares estimates. Based on the data in Table 4.3, the computer output shows that $b_1 = 1.758$, $b_2 = -0.352$, and $a = 2.529$. Consequently, the estimated regression equation is

$$Y_i = 2.529 + 1.758X_i - 0.352P_i \quad (4.12)$$

The estimated value of B_1 is 1.758, as contrasted with our earlier estimate of B , which was 1.504. In other words, a \$1 million increase in marketing expense results in an increase in estimated sales of 1.758 million units, as contrasted with

5. As in the case of simple regression, it is assumed that the mean value of e_i is zero and that the values of e_i are statistically independent (recall footnote 2).

TABLE 4.3

**Sales, Selling Expense, and Price, Miller Pharmaceutical Company,
Sample of Nine Years**

Selling Expense (Millions of Dollars)	Sales (Millions of Units)	Price (less \$10)
2	6	0
1	4	1
8	16	2
5	10	3
6	12	4
4	8	5
7	12	6
9	16	7
8	14	8

1.504 million units in the simple regression in equation (4.8). The reason these estimates differ is that the present estimate of the effect of marketing expense on sales holds constant the price, whereas the earlier estimate did not hold this factor constant. Since this factor affects sales, the earlier estimate is likely to be a biased estimate of the effect of marketing expense on sales.⁶

SOFTWARE PACKAGES AND COMPUTER PRINTOUTS

With few exceptions, regression analyses are carried out on computers, not by hand. Therefore, it is important that you know how to interpret computer printouts showing the results of regression calculations. Because there is a wide variety of “canned” programs for calculating regressions, no single format or list of items is printed out. However, the various kinds of printouts are sufficiently similar that it is worthwhile looking at two illustrations—Minitab and SAS—in some detail.

Figure 4.7 shows the Minitab printout from the multiple regression of the Miller Pharmaceutical Company’s sales (designated as C1) on its marketing expense (C2) and price (C3). According to this printout, the regression equation is

$$C1 = 2.529 + 1.758C2 - 0.352C3$$

The column headed “Coef” shows the estimated regression coefficient of each independent variable (called a “Predictor” on the printout). The intercept of the regression is the top figure in this vertical column (the figure in the horizontal row where the “Predictor” is “Constant”). The coefficient of determination (called R-sq)

6. Of course, this regression is supposed to be appropriate only when X_i and P_i vary in a certain limited range. If P_i is large and X_i is small, the regression would predict a negative value of sales, which obviously is inadmissible. But, as long as the regression is not used to make predictions for values of X_i and P_i outside the range of the data given in Table 4.3, this is no problem. For simplicity, we assume in equation (4.10) that the effect of price on the mean value of sales (holding marketing expense constant) can be regarded as linear in the relevant range. Alternatively, we could have assumed that it was quadratic or the constant-elasticity demand function discussed in Chapter 2 might have been used.

FIGURE 4.7

Minitab Printout of Results of Multiple Regression

MTB > regress c1 on 2 predictors in c2 and c3

The regression equation is
 $C1 = 2.53 + 1.76 C2 - 0.352 C3$

Predictor	Coef	Stdev	t-ratio	p
Constant	2.5294	0.2884	8.77	0.000
C2	1.75805	0.06937	25.34	0.000
C3	-0.35187	0.07064	-4.98	0.002

s = 0.3702 R-sq = 99.4% R-sq(adj) = 99.2%

Analysis of Variance

SOURCE	DF	SS	MS	F	p
Regression	2	144.067	72.033	525.72	0.000
Error	6	0.822	0.137		
Total	8	144.889			

SOURCE	DF	SEQ SS
C2	1	140.667
C3	1	3.399

is shown in the middle of the printout. For a multiple regression, the coefficient of determination is often called the *multiple coefficient of determination*.⁷

Figure 4.8 shows the SAS printout for the same regression. To find the intercept of the equation, obtain the figure (2.529431) in the horizontal row labeled “INTERCEP” that is in the vertical column called “Parameter Estimate.” To find the regression coefficient of marketing expense, obtain the figure (1.758049) in the horizontal row labeled “C2” that is in the vertical column called “Parameter Estimate.” To find the regression coefficient of price, obtain the figure (−0.351870) in the horizontal row labeled “C3” that is in the vertical column called “Parameter Estimate.” The multiple coefficient of determination is the figure (0.9943) to the right of “R-square.”

7. The positive square root of the multiple coefficient of determination is called the *multiple correlation coefficient*, denoted R . It too is sometimes used to measure how well a multiple regression equation fits the data.

The *unadjusted* multiple coefficient of determination— R -sq in Figure 4.7—can never decrease as another independent variable is added; a related measure without this property is the *adjusted* multiple coefficient of determination— R -sq (adj.) in Figure 4.7. The latter is often denoted \bar{R}^2 .

INTERPRETING THE OUTPUT OF STATISTICAL SOFTWARE

The following additional statistics are also of considerable importance: the standard error of estimate, the F statistic, and the t statistic. Each is discussed briefly next. For more detailed discussions of each, see any business statistics textbook.

The Standard Error of Estimate

A measure often used to indicate the accuracy of a regression model is the standard error of estimate, which is *a measure of the amount of scatter of individual observations about the regression line*. The standard error of estimate is denoted by “s” in the Minitab printout in Figure 4.7 and by “ROOT MSE” in the SAS printout in Figure 4.8. A comparison of these printouts shows that, in the Miller Pharma-

FIGURE 4.8

SAS Printout of Results of Multiple Regression

Dependent Variable: C1

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	2	144.06678	72.03339	525.718	0.0001
Error	6	0.82211	0.13702		
C Total	8	144.88889			
Root MSE	0.37016	R-square	0.9943		
Dep Mean	10.88889	Adj R-sq	0.9924		
C.V.	3.39944				

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > T
INTERCEP	1	2.529431	0.28842968	8.770	0.0001
C2	1	1.758049	0.06937127	25.343	0.0001
C3	1	-0.351870	0.07064425	-4.981	0.0025

ceutical multiple regression, the standard error is about 0.37 million units of sales. Of course, the answer is always the same, no matter which package we use.

To illustrate what the standard error of estimate measures, consider Figure 4.9. In panel A, the standard error of estimate is 1.5, which is much higher than in panel D, where it is 0.25. This is reflected in the much greater scatter in the points around the regression line in panel A than in panel D. As pointed out already, what the standard error of estimate measures is the amount of such scatter. Clearly, the amount of scatter decreases as we move from panel A to panel B to panel C to panel D. Similarly, the standard error of estimate decreases as we move from panel A to panel B to panel C to panel D.

The standard error of estimate is useful in constructing prediction intervals, that is, intervals within which there is a specified probability that the dependent variable will lie. If this probability is set at 0.95, a very approximate prediction interval is

$$\hat{Y} \pm 2s_e \quad (4.13)$$

where \hat{Y} is the predicted value of the dependent variable based on the sample regression and s_e is the standard error of estimate. For example, if the predicted value of the Miller Pharmaceutical Company's sales is 11 million units, the probability is about 0.95 that the firm's sales will be between 10.26 ($= 11 - 2 \times 0.37$) million units and 11.74 ($= 11 + 2 \times 0.37$) million units. However, it is important to note that equation (4.13) is a good approximation only if the independent variable is close to its mean; if this is not true, more complicated formulas must be used instead.⁸

8. The formula for the standard error of estimate is

$$\left[\sum_{i=1}^n (Y_i - \hat{Y}_i)^2 / (n - k - 1) \right]^{0.5}$$

where k is the number of independent variables.

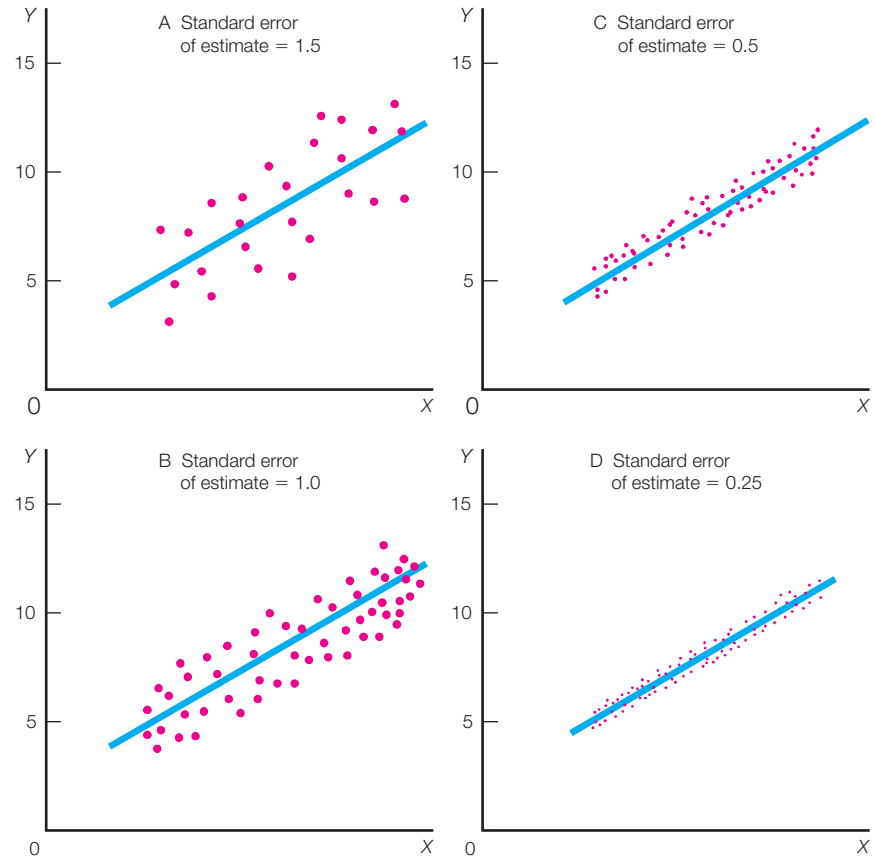
If the error term is normally distributed (see Appendix B for a description of the normal distribution), the exact prediction interval (with 0.95 probability) is

$$\hat{Y} \pm t_{0.025} s_e \left[\frac{n+1}{n} + \frac{(X^* - \bar{X})^2}{\sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2 / n} \right]^{0.5}$$

where $t_{0.025}$ is the value of a variable with the t distribution with $(n - 2)$ degrees of freedom that is exceeded with probability of 0.025, X^* is the value of the independent variable, and n is the sample size. (The t distribution is taken up in Appendix E.) This assumes that there is only one independent variable.

FIGURE 4.9

Four Regression Lines: Standard Error of Estimate Equals 1.5, 1.0, 0.5, and 0.25



The F Statistic

Frequently, a manager wants to know whether any of the independent variables really influences the dependent variable. In the case of the Miller Pharmaceutical Company, the marketing director may ask whether the data indicate that either marketing expense or price really influences the firm's sales. To answer such a question, one utilizes the F statistic, which is also included in the printout. The value of F is provided in the fifth horizontal row from the bottom of figures in the Minitab printout (Figure 4.7) and in the top horizontal row of figures in the SAS printout (Figure 4.8). Both printouts indicate that the value of F in the Miller Pharmaceutical case equals about 525.72.

Large values of F tend to imply that at least one of the independent variables has an effect on the dependent variable. Tables of the F distribution, a probability distribution named (or initialed) after the famous British statistician R. A. Fisher, are used to determine *the probability that an observed value of the F statistic could have arisen by chance, given that none of the independent variables has any effect on the dependent variable* (see Appendix E). This probability too is shown in the printout. It is denoted by “p” (immediately to the right of F) in the Minitab printout, and by “Prob > F” (immediately to the right of F VALUE) in the SAS printout. The value of this probability is 0.0001 (SAS) or 0.000 (Minitab); the difference is due to rounding.

Having this probability in hand, it is easy to answer the marketing director’s question. Clearly, the probability is extremely small—only about 1 in 10,000—that one could have obtained such a strong relationship between the dependent and independent variables sheerly by chance. Therefore, the evidence certainly suggests that marketing expense or price (or both) really influences the firm’s sales.

The t Statistic

Managers and analysts often are interested in whether a particular independent variable influences the dependent variable. For example, the president of the Miller Pharmaceutical Company may want to determine whether the amount allocated to marketing expense really affects the firm’s sales. As we know from equation (4.12), the least-squares estimate of B_1 is 1.758, which suggests that marketing expense has an effect on sales. But this least-squares estimate varies from one sample to another, and by chance it may be positive even if the true value of B_1 is zero.

To test whether the true value of B_1 is zero, we must look at the t statistic of B_1 , which is presented in the printout. For Minitab, recall that B_1 is the regression coefficient of C2, since the marketing expense is denoted by C2. Therefore, to find the t statistic for B_1 , we must locate the horizontal row of figures in the printout where the “Predictor” is C2 and obtain the figure in the vertical column called “t-ratio.” If SAS is used, find the horizontal row of figures where the “Variable” is C2 and obtain the figure in the vertical column called “T for H0: Parameter = 0.” If the error terms in the regression (that is, e_i) are normally distributed, the t statistic has a well-known probability distribution—the t distribution (see Appendix E).

All other things equal, the bigger is the value of the t statistic (in absolute terms), the smaller the probability that the true value of the regression coefficient in question really is zero. Based on the t distribution, it is possible to calculate *the probability, if the true value of the regression coefficient is zero, that the t statistic is as large (in absolute terms) as we observe*. This probability too is presented in the printout. For both Minitab and SAS, this probability is immediately to the right of the t statistic. For Minitab, it is in the vertical column labeled “p”; for SAS, it is in

Multicollinearity A condition where two or more independent variables are very highly correlated. Such a condition impacts the accuracy of the coefficient estimates.

9. Note that this is a *two-tailed test* of the hypothesis that marketing has no effect on sales. That is, it is a test of this hypothesis against the alternative hypothesis that the true regression coefficient of marketing is either positive or negative. In many cases, a *one-tailed test*—for example, in which the alternative hypothesis states that the true regression coefficient is positive only—may be more appropriate.

Frequently, a manager would like to obtain an interval estimate for the true value of a regression coefficient. In other words, he or she wants an interval that has a particular probability of including the true value of this regression coefficient. To find an interval that has a probability equal to $(1 - \alpha)$ of including this true value, you can calculate

$$b_1 \pm t_{\alpha/2} s_{b_1} \quad (4.14)$$

where s_{b_1} is the standard error of b_1 (in the horizontal row labeled “C2” and the vertical column labeled “Stdev” in the Minitab printout, or in the horizontal row labeled “C2” and the vertical column labeled “Standard Error” in the SAS printout) and where $t_{\alpha/2}$ is the $\alpha/2$ point on the t distribution with $(n - k - 1)$ degrees of freedom (see Appendix E). If α is set equal to 0.05, you obtain an interval that has a 95% probability of including B_1 . In the case of the Miller Pharmaceutical Company, since $B_1 = 1.758$, $s_{b_1} = 0.069$, and $t_{0.025} = 2.447$ it follows that a 95% confidence interval for B_1 is

$$1.758 \pm 2.447 (0.069)$$

or 1.589 to 1.927. For further discussion, see any business statistics textbook.

the vertical column labeled “Prob > T.” Regardless of whether Minitab or SAS is used, this probability is shown to be about 0.0001 (see Figures 4.7 and 4.8).

Given this probability, we can readily answer the question put forth by the president of the Miller Pharmaceutical Company. Recall that the president wanted to know whether the amount allocated to marketing really affects the sales. As the results show, it seems extremely likely the amount allocated to marketing really does affect sales. After all, according to the previous paragraph, the probability is only about 1 in 10,000 that chance alone would result in as large a t statistic (in absolute terms) as we found.⁹

MULTICOLLINEARITY

One important problem that can arise in multiple regression studies is **multicollinearity**, a situation in which two or more independent variables are very highly correlated. In the case of the Miller Pharmaceutical Company, suppose there had been a perfect linear relationship in the past between the firm’s marketing expense and its price. In a case of this sort, it is impossible to estimate the regression coefficients of both independent variables (X and P) because the data provide no information concerning the effect of one independent variable, holding the other independent variable constant. All that can be observed is the effect of both independent variables together, given that they both move together in the way they have in previous years.

Regression analysis estimates the effect of each independent variable by seeing how much effect this one independent variable has on the dependent variable when other independent variables are held constant. If two independent variables move together in a rigid, lockstep fashion, there is no way to tell how much effect each has separately; all we can observe is the effect of both combined. If there is good reason to believe that the independent variables will continue to move in lockstep in the future as they have in the past, multicollinearity does not prevent us from using regression analysis to predict the dependent variable. Since the two independent variables are perfectly correlated, one of them in effect stands for both and we therefore need use only one in the regression analysis. However, if the independent variables cannot be counted on to continue to move in lockstep, this procedure is dangerous, since it ignores the effect of the excluded independent variable.

In reality, managers seldom encounter cases in which independent variables are perfectly correlated, but they do encounter cases in which independent variables are so highly correlated that, although it is possible to estimate the regression coefficient of each variable, these regression coefficients cannot be estimated at all accurately. To cope with such situations, it sometimes is possible to alter the independent variables in such a way as to reduce multicollinearity. Suppose that a managerial economist wants to estimate a regression equation where the quantity demanded per year of a certain good is the dependent variable and the average

price of this good and disposable income of U.S. consumers are the independent variables. If disposable income is measured in money terms (that is, without adjustment for changes in the price level), there may be a high correlation between the independent variables. But if disposable income is measured in real terms (that is, with adjustment for changes in the price level), this correlation may be reduced considerably. Therefore, the managerial economist may decide to measure disposable income in real rather than money terms to reduce multicollinearity.

If techniques of this sort cannot reduce multicollinearity, there may be no alternative but to acquire new data that do not contain the high correlation among the independent variables.

SERIAL CORRELATION

In addition to multicollinearity, another issue that can occur in regression analysis is that the error terms (the values of e_i) are not independent; instead, they are serially correlated. For example, Figure 4.10 shows a case in which, if the error term in one period is positive, the error term in the next period is almost always positive. Similarly, if the error term in one period is negative, the error term in the next period almost always is negative. In such a situation, we say that the errors are *serially correlated* (or *autocorrelated*, which is another term for the same thing).¹⁰ Because this violates the assumptions underlying regression analysis, it is important that we be able to detect its occurrence. (Recall that regression analysis assumes that the values of e_i are independent.)

To see whether serial correlation is present in the error terms in a regression, we can use the Durbin-Watson test. Let \hat{e}_i be the difference between Y_i and \hat{Y}_i , the value of Y_i predicted by the sample regression. To apply the Durbin-Watson test, we (or in most cases, the computer) must calculate

$$d = \frac{\sum_{i=2}^n (\hat{e}_i - \hat{e}_{i-1})^2}{\sum_{i=1}^n \hat{e}_i^2} \quad (4.15)$$

Durbin and Watson provided tables that show whether d is so high or so low that the hypothesis that there is no serial correlation should be rejected. (Note that d is often called the Durbin-Watson statistic.)

Suppose we want to test this hypothesis against the alternative hypothesis that there is **positive serial correlation**. (Positive serial correlation would mean that e_i is directly related to e_{i-1} , as in Figure 4.10.) If so, we should reject the hypothesis of no serial correlation if $d < d_L$ and accept this hypothesis if $d > d_U$. If $d_L \leq d \leq d_U$, the test is inconclusive. The values of d_L and d_U are shown in Appendix E Table E-7. (Note that these values depend on the sample size n and on k , the number of independent variables in the regression.) On the other hand,

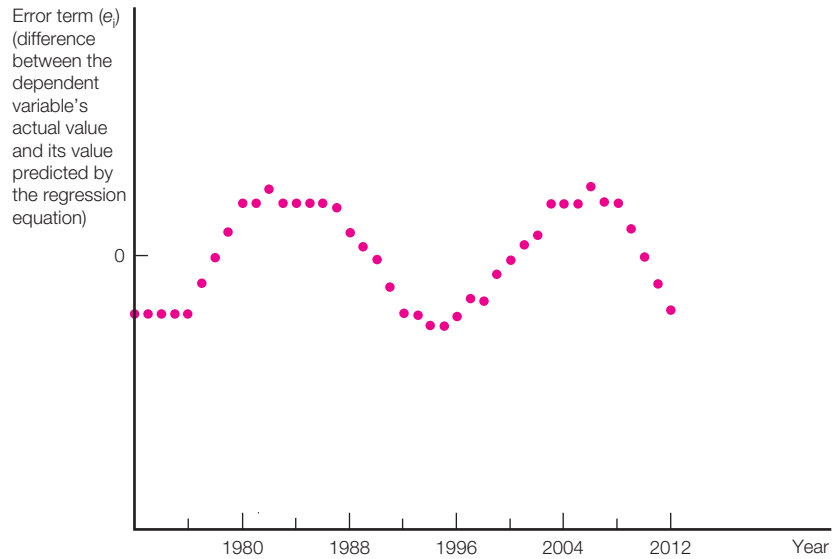
Positive serial correlation Occurs when error terms in regression analysis have the same sign as in the previous time period.

10. This is a case of positive serial correlation. (It is the sort of situation frequently encountered in managerial economics.) If the error term in one period tends to be positive (negative) and if the error term in the previous period is negative (positive), this is a case of negative serial correlation. More is said about this subsequently.

FIGURE 4.10

Serial Correlation of Error Terms

If the error term in one year is positive, the error term in the next year is almost always positive. If the error term in one year is negative, the error term in the next year is almost always negative.



Negative serial correlation

Occurs when error terms in regression analysis have the opposite sign as in the previous time period.

11. For a two-tailed test of both positive and negative serial correlation, reject the hypothesis of no serial correlation if $d < d_L$ or if $d > 4 - d_L$, and accept this hypothesis if $d_U < d < 4 - d_U$. Otherwise, the test is inconclusive. For a two-tailed test, the significance level is double the significance level shown in Appendix E Table E-7.

12. The use of first differences, while useful in some cases, is not always appropriate. For further discussion, see Johnston, *Econometric Methods*.

It is also important to avoid specification errors, which result when one or more significant explanatory variables is not included in the regression. If specification errors arise, the estimated regression coefficients may be biased and the regression equation may not predict very well. Also, problems can arise if the independent variables in a regression contain substantial measurement errors, since the regression coefficients of these variables often tend to be biased toward zero.

suppose the alternative hypothesis is that there is **negative serial correlation**. (Negative serial correlation means that e_i is inversely related to e_{i-1} .) If so, we should reject the hypothesis of no serial correlation if $d > 4 - d_L$ and accept this hypothesis if $d < 4 - d_U$. If $4 - d_U \leq d \leq 4 - d_L$, the test is inconclusive.”¹¹

One way to deal with the problem of serial correlation, if it exists, is to take first differences of all the independent and dependent variables in the regression. For example, in the case of the Miller Pharmaceutical Company, we might use the change in sales relative to the previous year (rather than the level of sales) as the dependent variable. And the change in marketing expense relative to the previous year (rather than the level of marketing expense) and the change in price relative to the previous year (rather than the level of price) might be used as the independent variables in the regression.¹²

FURTHER ANALYSIS OF THE RESIDUALS

In the previous section, we used \hat{e}_i (the difference between the actual value of Y_i and its value predicted by the sample regression) to test for serial correlation. Since it is a measure of the extent to which Y_i cannot be explained by the regression, \hat{e}_i

is often called the **residual** for the i th observation. Now we describe additional ways in which the residuals—that is, the values of \hat{e}_i —can be used to test whether the assumptions underlying regression analysis are met. We begin by plotting the value of each residual against the value of the independent variable. (For simplicity, we suppose only one independent variable.) That is, we plot \hat{e}_i against X_i , which is the independent variable.

Suppose that the plot is as shown in Figure 4.11. As you can see, the values of the residuals are much more variable when X_i is large than when it is small. In other words, the variation in \hat{e}_i increases as X_i increases. Since regression analysis assumes that *the variation in the error terms is the same, regardless of the value of the independent variable*, the plot in Figure 4.11 indicates that this assumption is violated. Two ways to remedy this situation are to use a weighted least-squares regression or to change the form of the dependent variable. For example, we might use $\log Y$ rather than Y as the dependent variable.¹³

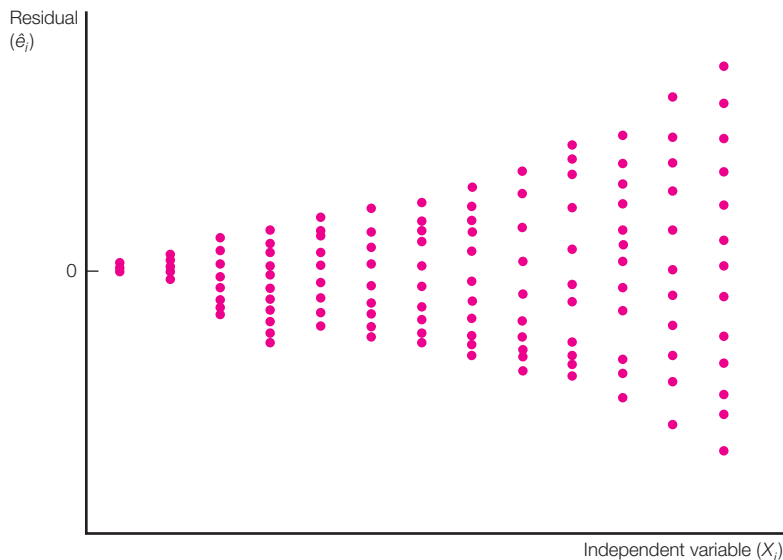
If the plot of \hat{e}_i against X_i looks like Figure 4.12, this is an indication that the relationship between the dependent and independent variables is not linear. When X is very low and very high, the linear regression *overestimates* the dependent variable, as shown by the fact that the residuals tend to be negative. When X is

Residual The difference between the presumed value of the dependent variable and the actual value of the dependent variable in any observation.

FIGURE 4.11

Residuals Indicating That the Variation in the Error Terms Is Not Constant

As you can see, the residuals vary less when X is small than when it is large.



13. For further details, see Johnston, *Econometric Methods*.

PROBLEM SOLVED: How Fed Economists Forecast Auto Output

Since purchases by the auto industry account for more than half of the rubber and lead consumed in this country as well as a major portion of the steel, aluminum, and a variety of other materials, it is obvious that many firms and government agencies, as well as the auto firms themselves, are interested in forecasting auto output. The Federal Reserve Bank of New York has published an article describing how the regression techniques described in this chapter have been used for this purpose. According to the author, Ethan Harris, the quantity of autos produced quarterly depends on five variables: (1) real disposable income, (2) the ratio of retail auto inventories to sales, (3) the average price of new cars (relative to the overall consumer price index), (4) the price level for nonauto durable goods, and (5) the prime rate (the interest rate banks charge their best customers).

The regression results follow. The probability that the t statistic for each of the regression coefficients is as large (in absolute terms) as it is here, if the true value of the regression coefficient is zero, is less than 0.01, except for the case of the nonauto price.

The value of the adjusted multiple coefficient of determination is 0.862, the standard error of estimate is 532, and the Durbin-Watson statistic (d) is 2.26. According to Harris, this regression equation has predicted auto output with a mean (absolute) error of about 6.9%.

(a) Would you expect the regression coefficient of the inventory-sales ratio to be negative? If so, why? (b) Can we be reasonably sure that the true value of the regression coefficient of the inventory-sales ratio is not zero? Why or why not? (c) Is there evidence of

positive serial correlation of the error terms? (d) Can we use this regression as an estimate of the demand curve for autos? Why or why not?

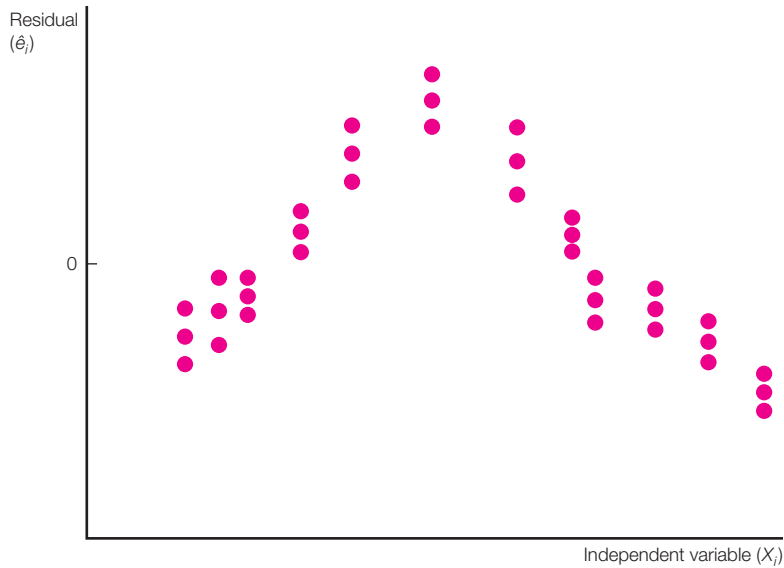
SOLUTION (a) Yes. If inventories are large relative to sales, one would expect auto firms to produce less than they would if inventories were small. (b) Yes. According to the preceding discussion, the probability that the t statistic for the regression coefficient of the inventory-sales ratio would be as great as 6.1 (in absolute terms) would be less than 0.01 if the true regression coefficient were zero. Hence, if this true regression coefficient were zero, it is exceedingly unlikely that the t statistic (in absolute terms) would equal its observed value or more. (c) No. Since the value of n is approximately 50 and $k = 5$, Appendix E Table E-7 shows that $d_L = 1.26$ and $d_U = 1.69$ if the significance level equals 0.025. The observed value of the Durbin-Watson statistic (2.26) is greater than d_U (1.69); this means that we should accept the hypothesis that there is no positive serial correlation. (d) No. One important indication that this is true is that the regression coefficient of the auto price is positive. Clearly, this regression equation cannot be used as an estimate of the demand curve for autos.

Variable	Regression coefficient	t statistic
Constant	-22,302	-4.5
Disposable income	12.9	6.6
Prime rate	-97.8	-3.2
Inventory-sales ratio	-19.9	-6.1
Auto price	230	5.0
Nonauto price	6.0	2.1

FIGURE 4.12

Residuals Indicating That the Relationship between the Dependent and Independent Variables Is Nonlinear, Not Linear

The residuals are negative when X is very small or very large and positive when X is of medium size.



of medium size, the linear regression underestimates the dependent variable, as shown by the fact that the residuals tend to be positive. It appears that a quadratic relationship fits the data better than a linear one. So, rather than assume that equation (4.2) holds, we should assume that

$$Y_i = A + B_1X_i + B_2X_i^2 + e_i$$

Using the multiple regression techniques described previously, the values of A , B_1 , and B_2 can be estimated.

SUMMARY

1. An identification problem can occur if price in various periods is plotted against quantity demanded and the resulting relationship is used to estimate the demand curve. Because nonprice variables are not held constant, the demand curve may have shifted over time. Nonetheless, sophisticated econometric methods can be used to estimate the demand function. Also, market experiments and consumer

interviews may be of value. For example, managers sometimes vary price from one city or region to another, to see what the effects are on quantity demanded.

2. Regression analysis is useful in estimating demand functions and other economic relationships. The regression line shows the average relationship between the dependent variable and the independent variable. The method of least squares is the standard technique used to fit a regression line to a set of data. If the regression line is $\hat{Y} = a + bX$ and if a and b are calculated by least squares,

$$b = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

and

$$a = \bar{Y} - b\bar{X}$$

This value of b is often called the *estimated regression coefficient*.

3. Whereas a simple regression includes only one independent variable, a multiple regression includes more than one independent variable. An advantage of multiple regression over a simple regression is that you frequently can predict the dependent variable more accurately if more than one independent variable is used. Also, if the dependent variable is influenced by more than one independent variable, a simple regression of the dependent variable on a single independent variable may result in a biased estimate of the effect of this independent variable on the dependent variable.

4. The first step in multiple regression analysis is to identify the independent variables and specify the mathematical form of the equation relating the mean value of the dependent variable to the independent variables. For example, if Y is the dependent variable and X and P are identified as the independent variables, one might specify that

$$Y_i = A + B_1X_i + B_2P_i + e_i$$

where e_i is an error term. To estimate B_1 and B_2 (called the *true regression coefficients* of X and P) as well as A (the intercept of this true regression equation), we use the values that minimize the sum of squared deviations of Y_i from \hat{Y}_i , the value of the dependent variable predicted by the estimated regression equation.

5. In a simple regression, the coefficient of determination is used to measure the closeness of fit of the regression line. In a multiple regression, the multiple coefficient of determination, R^2 , plays the same role. The closer R^2 is to 0, the poorer the fit; the closer it is to 1, the better the fit.

6. The F statistic is used to test whether any of the independent variables has an effect on the dependent variable. The standard error of estimate can help to indicate how well a regression model can predict the dependent variable. The

t statistic for the regression coefficient of each independent variable is used to test whether this independent variable has any effect on the dependent variable.

7. A difficult problem that can occur in multiple regression is multicollinearity, a situation in which two or more of the independent variables are highly correlated. If multicollinearity exists, it may be impossible to estimate accurately the effect of particular independent variables on the dependent variable. Another frequently encountered problem arises when the error terms in a regression are serially correlated. The Durbin-Watson test can be carried out to determine whether this problem exists. Plots of the residuals can help to detect cases in which the variation of the error terms is not constant or where the relationship is nonlinear not linear.

PROBLEMS

wwwnorton.com/studyspace 

1. The Klein Corporation's marketing department, using regression analysis, estimates the firm's demand function, the result being

$$Q = -104 - 2.1P + 3.2I + 1.5A + 1.6Z$$

$$R^2 = 0.89$$

Standard error of estimate = 108

where Q is the quantity demanded of the firm's product (in tons), P is the price of the firm's product (in dollars per ton), I is per capita income (in dollars), A is the firm's advertising expenditure (in thousands of dollars), and Z is the price (in dollars) of a competing product. The regression is based on 200 observations.

- a. According to the statistical software, the probability is 0.005 that the t statistic for the regression coefficient of A would be as large (in absolute terms) as it is in this case if in fact A has no effect on Q . Interpret this result.
 - b. If $I = 5,000$, $A = 20$, and $Z = 1,000$, what is the Klein Corporation's demand curve?
 - c. If $P = 500$ (and the conditions in part b hold), estimate the quantity demanded of the Klein Corporation's product.
 - d. How well does this regression equation fit the data?
2. Since all the Hawkins Company's costs (other than advertising) are essentially fixed costs, managers want to maximize total revenue (net of advertising expenses). According to a regression analysis (based on 124 observations) carried out by managers,

$$Q = -23 - 4.1P + 4.2I + 3.1A$$

where Q is the quantity demanded of the firm's product (in dozens), P is the price of the firm's product (in dollars per dozen), I is per capita income (in dollars), and A is advertising expenditure (in dollars).

- a. If the price of the product is \$10 per dozen, should managers increase advertising?

- b. If the advertising budget is fixed at \$10,000, and per capita income equals \$8,000, what is the firm's marginal revenue curve?
 - c. If the advertising budget is fixed at \$10,000, and per capita income equals \$8,000, what price should managers charge?
3. The 2012 sales and profits of seven clothing companies were as follows:

Firm	Sales (\$ billions)	Profit (\$ billions)
Maxx	5.7	0.27
Bleu	6.7	0.12
Golden	0.2	0.00
Triex	0.6	0.04
Chateau	3.8	0.05
L & T	12.5	0.46
Eastview	0.5	0.00

- a. Calculate the sample regression line, where profit is the dependent variable and sales is the independent variable.
 - b. Estimate the 2012 average profit of a clothing firm with 2012 sales of \$0.2 billion.
 - c. Can this regression line be used to predict a clothing firm's profit in 2026? Explain.
4. The Cherry Manufacturing Company's chief engineer examines a random sample of 10 spot welds of steel. In each case, the shear strength of the weld and the diameter of the weld are determined, the results being as follows:

Shear Strength (Pounds)	Weld Diameter (Thousandths of an Inch)
680	190
800	200
780	209
885	215
975	215
1,025	215
1,100	230
1,030	250
1,175	265
1,300	250

- a. Does the relationship between these two variables seem to be direct or inverse? Does this accord with common sense? Why or why not? Does the relationship seem to be linear?

- b. Calculate the least-squares regression of shear strength on weld diameter.
- c. Plot the regression line. Use this regression line to predict the average shear strength of a weld $\frac{1}{5}$ inch in diameter. Use the regression line to predict the average shear strength of a weld $\frac{1}{4}$ inch in diameter.
5. The Kramer Corporation's marketing manager calculates a regression, where the quantity demanded of the firm's product (designated as "C1") is the dependent variable and the price of the product (designated as "C2") and consumers' disposable income (designated as "C3") are independent variables. The Minitab printout for this regression follows

MTB > regress c1 on 2 predictors in c2 and c3
 The regression equation is
 $C1 = 40.8 - 1.02 C2 + 0.00667 C3$

Predictor	Coef	Stdev	t-ratio	p
Constant	40.833	1.112	36.74	0.000
C2	-1.02500	0.06807	-15.06	0.000
C3	0.006667	0.005558	1.20	0.244

S=1.361 R-sq=91.62% R-sq(adj)=90.8%
 Analysis of variance

SOURCE	DF	SS	MS	F	p
Regression	2	422.92	211.46	114.11	0.000
Error	21	38.92	1.85		
Total	23	461.83			

SOURCE	DF	SEQ SS
C2	1	420.25
C3	1	2.67

- a. What is the intercept of the regression?
- b. What is the estimated regression coefficient of the product's price?
- c. What is the estimated regression coefficient of disposable income?
- d. What is the multiple coefficient of determination?
- e. What is the standard error of estimate?
- f. What is the probability that the observed value of the F statistic could arise by chance, given that neither of the independent variables has any effect on the dependent variable?
- g. What is the probability, if the true value of the regression coefficient of price is zero, that the t statistic is as large (in absolute terms) as we observe?
- h. What is the probability, if the true value of the regression coefficient of disposable income is zero, that the t statistic is as large (in absolute terms) as we observe?
- i. Describe briefly what this regression means.
6. Railroad executives must understand how the costs incurred in a freight yard are related to the output of the yard. The two most important services performed by a yard are switching and delivery, and it seems reasonable to use the number of cuts switched and the number of cars delivered during a particular period as a measure of output. (A cut is a group of cars that rolls as a unit onto

the same classification track; it is often used as a unit of switching output.) A study of one of the nation's largest railroads assumed that

$$C_i = A + B_1S_i + B_2D_i + e_i$$

where C_i is the cost incurred in this freight yard on the i th day, S_i is the number of cuts switched in this yard on the i th day, D_i is the number of cars delivered in this yard on the i th day, and e_i is an error term. Data were obtained regarding C_i , S_i , and D_i for 61 days. On the basis of the procedures described in this chapter, these data were used to obtain estimates of A , B_1 , and B_2 . The resulting regression equation was

$$\hat{C}_i = 4,914 + 0.42S_i + 2.44D_i$$

where \hat{C}_i is the cost (in dollars) predicted by the regression equation for the i th day.

- a. If you were asked to evaluate this study, what steps would you take to determine whether the principal assumptions underlying regression analysis were met?
 - b. If you were satisfied that the underlying assumptions were met, of what use might this regression equation be to the railroad? Be specific.
 - c. Before using the study's regression equation, what additional statistics would you like to have? Why?
 - d. If the Durbin-Watson statistic equals 2.11, is there evidence of serial correlation in the residuals?
7. Mary Palmquist, a Wall Street securities analyst, wants to determine the relationship between Chile's gross domestic product (GDP) and the profits (after taxes) of the Carlton Company. She obtains the following data concerning each variable

Year	Gross Domestic Product (Billions of Dollars)	Carlton's Profits (Millions of Dollars)
2001	688	355
2002	753	339
2003	796	361
2004	868	357
2005	936	278
2006	982	363
2007	1,063	510
2008	1,171	573
2009	1,306	661
2010	1,407	705
2011	1,529	688
2012	1,706	931

- a. What are the least-squares estimates of the intercept and slope of the true regression line, where Carlton's profits are the dependent variable and GDP is the independent variable?
 - b. On the average, what effect does a \$1 increase in gross domestic product seem to have on the profits of Carlton?
 - c. If Ms. Palmquist feels that next year's GDP will be \$2 trillion, what forecast of Carlton's profits will she make on the basis of the regression?
 - d. What is the coefficient of determination between the nation's gross domestic product and Carlton's profits?
 - e. Do the results obtained in previous parts of this problem prove that changes in Carlton's profits are caused by changes in the gross domestic product? Can we be sure that Carlton's profit is a linear function of the GDP? What other kinds of functions might be as good or better?
 - f. If you were the financial analyst, would you feel that this regression line was an adequate model to forecast Carlton's profits? Why or why not?
8. In the manufacture of cloth, the weft packages should not disintegrate unduly during weaving. A direct measure of the tendency to disintegrate exists, but it is laborious and uneconomical to carry out. In addition, there are indirect measures based on laboratory tests. Managers of the Brockway Textile Company want to determine the extent to which one of these indirect measures is correlated with the direct measure. If the correlation is high enough, Brockway managers believe that they may be able to use the indirect measure instead of the direct measure.

An experiment was carried out in which both the direct and indirect measures of the tendency to disintegrate were calculated for 18 lots of packages. The results follow

Lot	Measure		Lot	Measure	
	Direct	Indirect		Direct	Indirect
1	31	6.2	10	0	10.7
2	31	6.2	11	35	4.1
3	21	10.1	12	63	3.5
4	21	8.4	13	10	5.0
5	57	2.9	14	51	4.5
6	80	2.9	15	24	9.5
7	35	7.4	16	15	8.5
8	10	7.3	17	80	2.6
9	0	11.1	18	90	2.9

- a. What is the coefficient of determination between the two measures?

- b. What linear regression line would you use to predict the value of the direct measure on the basis of knowledge of the indirect measure?
 - c. On the basis of your findings, write a brief report indicating the factors to be weighed in deciding whether to substitute the indirect measure for the direct measure.
9. The Kingston Company hires a consultant to estimate the demand function for its product. Using regression analysis, the consultant estimates the demand function to be

$$\log Q = 2.01 - 0.148 \log P + 0.258 \log Z$$

where Q is the quantity demanded (in tons) of Kingston's product, P is the price (in dollars per ton) of Kingston's product, and Z is the price (in dollars per ton) of a rival product.

- a. Calculate the price elasticity of demand for Kingston's product.
 - b. Calculate the cross elasticity of demand between Kingston's product and the rival product.
 - c. According to the consultant, $\bar{R}^2 = 0.98$ and the standard error of estimate is 0.001. If the number of observations is 94, comment on the goodness of fit of the regression.
10. Managers of the New Hope and Ivyland Short Line Railroad conducted an experiment in which they reduced fares by about 28% for approximately a year to estimate the price elasticity of demand. This large fare reduction resulted in essentially no change in the railroad's revenues.
- a. What problems exist in carrying out an experiment of this sort?
 - b. Taken at face value, what seemed to be the price elasticity of demand?
11. Because of a shift in consumer tastes, the market demand curve for high-quality red wine has shifted steadily to the right. If the market supply curve has remained fixed (and is upward sloping to the right), there has been an increase over time in both the price of such wine and in the quantity sold.
- a. If one were to plot price against quantity sold, would the resulting relationship approximate the market demand curve?
 - b. If not, what would this relationship approximate?
12. Managers of the Brennan Company used regression analysis to obtain the following estimate of the demand function for their product

$$\log Q = 2 - 1.2 \log P + 1.5 \log I$$

where Q is quantity demanded, P is price, and I is consumers' disposable income.

- a. Brennan's president is considering a 5% price reduction. He argues that these results indicate that such action will result in a 6% increase in the number of units sold by the firm. Do you agree? Why or why not?

- b. The firm's treasurer calculates that the probability that the t statistic of $\log P$ is as large (in absolute value) as it is, given that $\log P$ has no real effect on $\log Q$, is about 0.5. He says the estimate of the price elasticity is unreliable. Do you agree? Why or why not?
- c. How can managers obtain a more accurate estimate of the price elasticity of demand?

APPENDIX: THE COEFFICIENT OF DETERMINATION AND THE CONCEPT OF EXPLAINED VARIATION

In this appendix, we provide a fuller explanation of what the coefficient of determination is and how it can be interpreted. To begin with, we must discuss the concept of variation, which refers to a sum of squared deviations. The total variation in the dependent variable Y equals

$$\sum_{i=1}^n (Y_i - \bar{Y})^2 \quad (4.16)$$

In other words, the total variation equals the sum of the squared deviations of Y from its mean.

To measure how well a regression line fits the data, we divide the total variation in the dependent variable into two parts: the variation that *can* be explained by the regression line and the variation that *cannot* be explained by the regression line. To divide the total variation in this way, we must note that, for the i th observation,

$$(Y_i - \bar{Y}) = (Y_i - \hat{Y}_i) + (\hat{Y}_i - \bar{Y}) \quad (4.17)$$

where \hat{Y}_i is the value of Y_i that would be predicted on the basis of the regression line. In other words, as shown in Figure 4.13, the discrepancy between Y_i and the mean value of Y can be split into two parts: the discrepancy between Y_i and the point on the regression line directly below (or above) Y_i and the discrepancy between the point on the regression line directly below (or above) Y_i and \bar{Y} .

It can be shown that¹⁴

$$\sum_{i=1}^n (Y_i - \bar{Y})^2 = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 + \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2 \quad (4.18)$$

The term on the left-hand side of this equation shows the total variation in the dependent variable. The first term on the right-hand side measures the *variation in the dependent variable not explained by the regression*. This is a reasonable interpretation of this term, since it is the sum of squared deviations of the actual observations from the regression line. Clearly, the larger is the value of this term, the poorer the regression equation fits the data. The second term on the right-hand side of the equation measures the *variation in the dependent variable explained by*

14. To derive this result, we square both sides of equation (4.17) and sum the result over all values of i . We find that

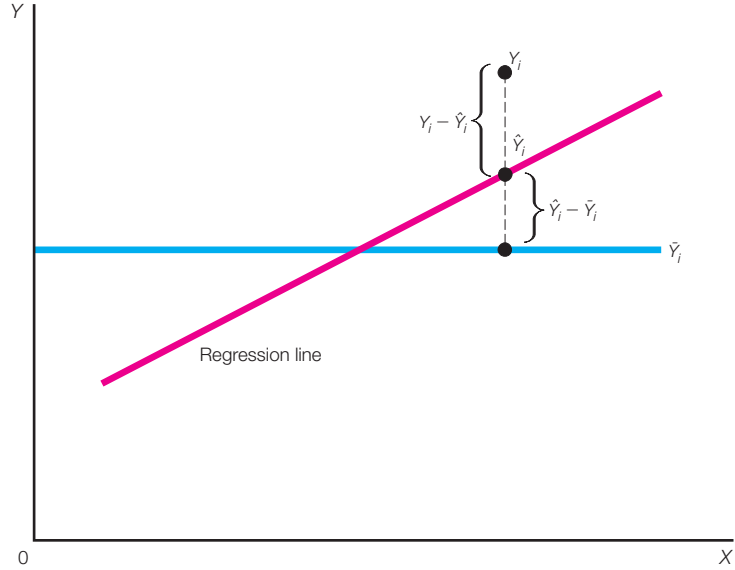
$$\begin{aligned} \sum_{i=1}^n (Y_i - \bar{Y})^2 &= \sum_{i=1}^n [(Y_i - \hat{Y}_i) + (\hat{Y}_i - \bar{Y})]^2 \\ &= \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 + \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2 \\ &\quad + 2 \sum_{i=1}^n (Y_i - \hat{Y}_i)(\hat{Y}_i - \bar{Y}) \end{aligned}$$

The last term on the right-hand side equals zero, so equation (4.18) follows.

FIGURE 4.13

Division of $(Y_i - \bar{Y})$ into Two Parts: $(Y_i - \hat{Y}_i)$ and $(\hat{Y}_i - \bar{Y})$

This division is carried out to measure how well the regression line fits the data.



the regression. This is a reasonable interpretation of this term, since it shows how much the dependent variable is expected to vary on the basis of the regression alone.

Coefficient of determination The proportion of the total variation in the dependent variable explained by the regression line.

To measure the closeness of fit of a simple regression line, we use the **coefficient of determination**, which equals

$$1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (4.19)$$

In other words, the coefficient of determination equals

$$\begin{aligned} 1 - \frac{\text{variation not explained by regression}}{\text{total variation}} \\ = \frac{\text{variation explained by regression}}{\text{total variation}} \end{aligned} \quad (4.20)$$

Clearly, the coefficient of determination is a reasonable measure of the closeness of fit of the regression line, since it equals the proportion of the total variation in the

dependent variable explained by the regression line. The closer it is to 1, the better the fit; the closer it is to 0, the poorer the fit.

When a multiple regression is calculated, the multiple coefficient of determination is used to measure the goodness of fit of the regression. The multiple coefficient of determination is defined as

$$R^2 = 1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (4.21)$$

where \hat{Y}_i is the value of the dependent variable that is predicted from the regression equation. So, as in the case of the simple coefficient of determination covered earlier,

$$R^2 = \frac{\text{variation explained by regression}}{\text{total variation}} \quad (4.22)$$

This means that R^2 *measures the proportion of the total variation in the dependent variable explained by the regression equation.*

PART 3

PRODUCTION AND COST

CHAPTER 5

LEARNING OBJECTIVES

The Production Function with One Variable Input

The Law of Diminishing Marginal Returns

The Production Function with Two Variable Inputs

Isoquants

The Marginal Rate of Technical Substitution

The Optimal Combination of Inputs

Corner Solutions

Returns to Scale

The Output Elasticity

Estimations of Production Functions

Summary

Problems

Appendix: Lagrangian Multipliers and Optimal Input Combinations

PRODUCTION THEORY

Once managers determine the demand for the firm's product or service, their job is far from over. Now they must choose the optimal method to produce. Managers need to be as efficient as possible. Resources are costly, and using them wisely is the hallmark of good managers. Efficiency requires an understanding of the production process. Simply stated, a production process explains how scarce resources (inputs) are used to produce a good or service (output). The production function precisely specifies the relationship between inputs and outputs.

Production issues are not confined to the physical transformation of inputs into outputs. In business, production involves all activities associated with providing goods and services, such as employment practices, acquisition of capital resources, and product distribution. Today, at firms like investment banks and consulting practices, managers are concerned with efficiently producing intellectual resources.

Understanding the production process is fundamental to gaining insight into cost analysis. Control of costs, along with an understanding of demand, is required for managers to optimize profit. But costs evolve from the production process. Managers cannot understand their firm's cost structure unless they understand the production process.

STRATEGY SESSION: The Yankees' Deal for Alex Rodriguez

In February 2004 the New York Yankees assumed the richest contract in sports by trading players to the Texas Rangers for shortstop Alex Rodriguez. Rodriguez was working under a 10-year, \$252 million contract that he signed with the Rangers in 2000. As soon as the deal was announced, commentators speculated about whether the contract was economically sensible. In fact, when we look at the underlying economics, the Yankees appear to have gotten the best player in baseball at a relatively bargain price.

First, the Yankees received \$67 million in cash from the Rangers—the largest sum to trade hands in the history of baseball. This reduced the Yankees' liability to Rodriguez to roughly \$112 million. The expected payouts to the players the Yankees got rid of in the trade were approximately \$13.3 million. Rodriguez was to receive \$15 million in 2004, but he agreed to defer \$1 million. Hence the added cost of Rodriguez to the Yankees' payroll in 2004 was roughly \$750,000.

For future years, the payouts to Rodriguez would be the following: \$15 million for the next three seasons; \$16 million in 2007–2008; \$17 million in 2009; and \$18 million in 2010. Rodriguez would defer \$1 million for the first four years at no interest and receive the \$4 million in 2011.

Because of their high payroll costs, the Yankees must pay a luxury tax (which is split among the other

baseball teams). In 2004 the Yankees estimated their payroll would equal \$190 million (this was higher than the 10 lowest-payroll teams). A payroll of \$190 million would cost the Yankees \$21 million in luxury taxes.

The total revenue for the Yankees was estimated to be \$330 million in 2004. Approximately \$110 million was generated from 3.5 million paying customers (ticket prices increased by an average of 10% for the 2004 season). The Yankees received \$60 million from the YES Network, \$10 million from WCBS radio, and over \$30 million from national television, licensing, and sponsorship revenues. The team also received revenue from local sponsorships and game concessions.

The Yankees expected the addition of Alex Rodriguez to increase attendance even in the face of increasing ticket prices (after the trade was announced, the ticket office was swamped with ticket requests). These additional fans also presumably would spend more at the games on snacks, drinks, and merchandise.

After all the factors are considered, most experts believe the marginal benefit to the Yankees was greater than the additional costs for Rodriguez's contract.

Source: "Sports Business: Steinbrenner Has Got It, and He Loves to Flaunt It," *New York Times*, February 17, 2004.

THE PRODUCTION FUNCTION WITH ONE VARIABLE INPUT

The production function is a table, a graph, or an equation showing the maximum product output achieved from any specified set of inputs. The function summarizes the characteristics of existing technology at a given time; it shows the technological constraints managers face. Any manager should want to use the most efficient process known. So we assume managers presuppose technical efficiency. Unfortunately many managers view processes as static. Production is dynamic: Methods,

designs, and factor costs change. Changes beget changes and may require different input mixes.

Say a process uses two inputs. If X_1 is the level of the first input and X_2 is the level of the second input, the production function is

$$Q = f(X_1, X_2) \quad (5.1)$$

where Q is the firm's output rate.

Cognitively, the simplest case has one input whose quantity is fixed and one input whose quantity is variable. Fixed inputs cannot be changed in the short run. To be sure, economists assume the time needed to change an asset is the beginning of what is called the long term. Fixed inputs often require capital (buildings, machinery, land). Variable inputs can be changed in the short run; labor is an example. In the long run, all inputs are variable.

John Thomas is an entrepreneur who currently owns five CNC machine tools. He works as a contractor in the airplane industry. He wants to know the effect

TABLE 5.1

Output of Metal Parts When Various Amounts of Labor Are Applied to Five Machine Tools, Thomas Machine Company

Amount of Labor (L)	Amount of Capital (Number of Machines)	Output of Parts (Q , hundreds per year)
0	5	0
1	5	49
2	5	132
3	5	243
4	5	376
5	5	525
6	5	684
6.67	5	792.59
7	5	847
8	5	1,008
9	5	1,161
10	5	1,300
11	5	1,419
12	5	1,512
13	5	1,573
14	5	1,596
15	5	1,575

on annual output if he were to hire various numbers of machinists. (Please note that the following output numbers are expressed in hundreds.) Thomas estimates that one machinist produces 49 parts per year. Thomas can produce more parts by hiring more workers, as we see in Table 5.1. This table represents a production function for Thomas Machine Company when five machine tools are used. More visually, the curve in Figure 5.1 presents exactly the same results. In fact, the

FIGURE 5.1

Relationship between Total Output and Amount of Labor Used on Five Machine Tools, Thomas Machine Company

Total output increases as labor increases at an increasing rate (up to 6.67 units of labor) and increases at a decreasing rate (until slightly more than 14 units of labor). Thereafter, output decreases as more units of labor are deployed. Managers will never willfully deploy labor in the latter circumstance. The production function shows the relationship between output (in this case number of parts produced) and input (in this case units of labor).

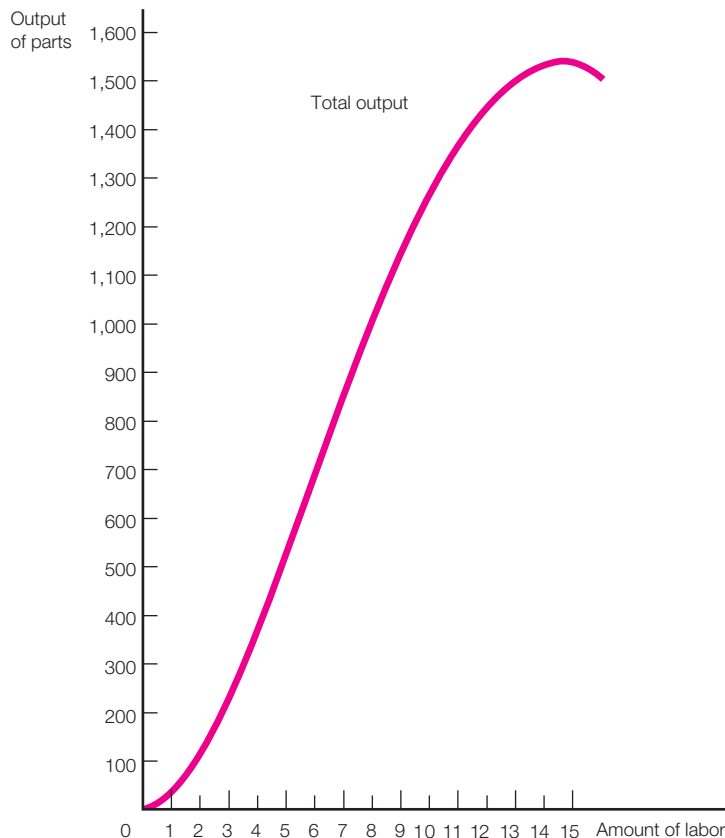


TABLE 5.2

Average and Marginal Products of Labor, Thomas Machine Company

Amount of Labor (Units)	Amount of Capital (Number of Machines)	Output of Parts (Q , Hundreds of Parts)	Average Product (Q/L)	Marginal Product ($\Delta Q/\Delta L$) ^a	Marginal Product (dQ/dL) ^a
0	5	0	—	—	—
1	5	49	49	49	67
2	5	132	66	83	98
3	5	243	81	111	123
4	5	376	94	133	142
5	5	525	105	149	155
6	5	684	114	159	162
6.67	5	792.59	118.89	162.89	163.33
7	5	847	121	163	163
8	5	1,008	126	161	158
9	5	1,161	129	153	147
10	5	1,300	130	139	130
11	5	1,419	129	119	107
12	5	1,512	126	93	78
13	5	1,573	121	61	43
14	5	1,596	114	23	2
15	5	1,575	105	−19	−45

^aThe figures in the $\Delta Q/\Delta L$ column pertain to the interval between the indicated amount of labor and one unit less than the indicated amount of labor. The figures in the dQ/dL column are the continuous marginal product—that is, $dQ/dL = MP_L = 30 + 40L - 3L^2$.

numbers in Table 5.1 (and Table 5.2) are derived from the production function equal to $Q = 30L + 20L^2 - L^3$. L is equal to the number of machinists.

We can think of the production function as giving insights into the manager's technology use. Thomas is clearly interested in knowing how output changes as the number of machinists varies. One common measure used by many managers is *output per worker*. This measure is what economists call **average product (AP)**. Because we are varying machinists, this is output per worker or

$$AP = \frac{Q}{X_1}, \text{ holding } X_2 \text{ constant}$$

Average product tells Thomas how many units of output, on average, each machinist is responsible for. If he wants a better metric to estimate the efficiency of each

Average product (AP) Common measuring device for estimating the units of output, on average per input unit.

worker, he should use what economists call the **marginal product (MP)**. The input's MP is equal to the incremental change in output created by a small change in input

$$MP = \frac{\Delta Q}{\Delta X_1}, \text{ holding } X_2 \text{ constant}$$

For machinists, the marginal product represents the impact on output of a unit change in machinists. If Thomas adds a machinist, the question is, "How many more units did we produce because I hired this last machinist?" If he must let one go, it is, "How many fewer units did we produce because I let this machinist go?" The marginal product is what Thomas wants to measure.

We calculate the average product and marginal product of labor, based on our estimated $Q = 30L + 20L^2 - L^3$. Both vary, of course, as we allocate the machinists to our five machines. If $Q(L)$ is total output with L units of labor per year, the average product is $Q(L)/L$. The marginal product of labor MP , when between L and $(L - 1)$ units of labor per year, is $Q(L) - Q(L - 1)$. From Table 5.2 we see that the average product of our first machinist is 49 parts and the marginal product is 83 parts per machinist between the first and second hires. Results for other machinist hires are shown in Table 5.2.

Marginal product (MP) Metric for estimating the efficiency of each input in which the input's MP is equal to the incremental change in output created by a small change in the input.

QUANT OPTION

More precisely, the marginal product of an input is the derivative of output with regard to the quantity of the input. That is, if Q is the output and x is the quantity of the input, the marginal product of the input equals dQ/dx if the quantities of all other inputs are fixed.

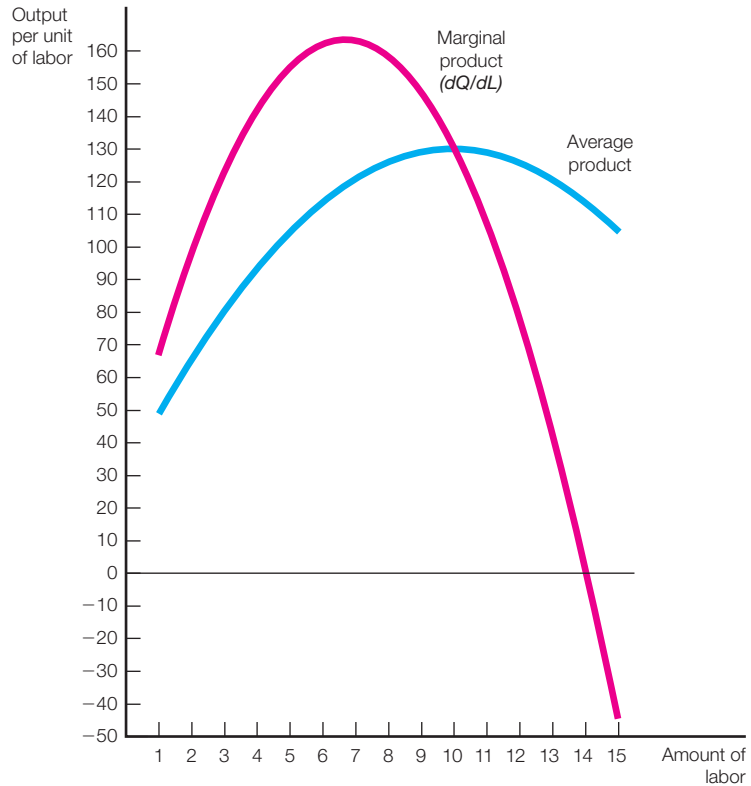
The average and marginal products of machinists are shown in Figure 5.2; the numbers are derived from Table 5.1. The curve is representative of most production processes. The average product of machinists (with five machines) reaches a maximum (at $L = 10$ and $Q/L = 130$), then falls. The marginal product of labor follows a similar pattern: It initially increases, reaches a maximum (at $L = 6.67$ and marginal product = 163.33), then falls. This, too, is typical of most production processes. Figure 5.2 shows that the marginal product equals the average product when the latter reaches a maximum; that is, $MP = AP = 130$ when $L = 10$.

We use two definitions of marginal product in Table 5.2. The first ($\Delta Q/\Delta L$) assumes that Thomas employs labor in discrete units, as in a machinist or a machinist-hour. This may be due to employment laws or negotiated contracts with

FIGURE 5.2

Average and Marginal Product Curves for Labor

Marginal product exceeds average product when the latter is increasing and is less than average product when the latter is decreasing. (Output per unit of labor is measured in hundreds of parts.)



labor. The second (dQ/dL) assumes that Thomas can employ labor continuously, as in 1.25 workers or 1.33 workers. This could be achieved by using part-time workers or workers who work more or less time than in a standard day's work.

We need to understand why MP equals AP when AP is maximized. A simple intuitive frame may help. Assume your professor is grading papers. She is keeping a running average of test scores by calculating a new average after each paper she corrects. After grading three papers her average is 86. If the next (marginal) test score is higher than 86, the average must increase. If it is lower than 86, the average must decrease. This is a natural law of mathematics. So as long as MP is greater than AP , AP must be increasing. When MP is less than AP , AP must be decreasing. MP intersects with AP when AP is at a maximum.

QUANT OPTION

Enough chat; let's get serious. If x is the input with Q as the output, then AP is Q/x and MP is dQ/dx

$$\begin{aligned}\frac{d(Q/x)}{dx} &= \frac{x(dQ/dx) - Q(dx/dx)}{x^2} \\ &= \frac{1}{x} \left(\frac{dQ}{dx} - \frac{Q}{x} \right)\end{aligned}$$

When the average product is at a maximum, $d(Q/x)/dx$ equals zero. Therefore,

$$\frac{d(Q/x)}{dx} = \frac{1}{x} \left(\frac{dQ}{dx} - \frac{Q}{x} \right) = 0$$

And hence

$$MP = \frac{dQ}{dx} = \frac{Q}{x} = AP$$

THE LAW OF DIMINISHING MARGINAL RETURNS

The law of diminishing returns is a well-known constraint in managerial economics and a good one for managers to understand. It teaches managers to remain in balance. For most production processes, if managers add equal increments of an input while holding other input levels constant, the incremental gains (MP) to output get smaller, and if pushed to the extreme, are counterproductive. It is not hard to see why diminishing marginal returns are found in most production functions.

We see in Table 5.2 that if Thomas hires an eighth machinist, marginal product will decrease. Why? Because Thomas has only five machines. As more machinists are hired, they will have to ration machines, or new hires will be assigned to less important tasks.

Choosing the optimal input bundle is not an easy managerial task. Managers cannot hold all inputs but one constant; and they cannot expect that adding more units will always result in large increases in output. It is not as simple as that, as we will see later.

Law of diminishing returns A well-known occurrence where when managers add equal increments of an input while holding other input levels constant, the incremental gains to output eventually get smaller.

THE PRODUCTION FUNCTION WITH TWO VARIABLE INPUTS

Now we want to complicate John Thomas' world. With a longer time horizon, the formerly fixed input of five CNC machines becomes variable. Table 5.3 shows the extra input combinations to consider. Though Thomas will have to consider more choices, the process is similar to that of the one-variable input case.

TABLE 5.3

Production Function, Two Variable Inputs, Thomas Machine Company

Amount of Labor (Units)	Quantity of Machine Tools (Hundreds of Parts Produced per Year)			
	3	4	5	6
1	5	11	18	24
2	14	30	50	72
3	22	60	80	99
4	30	81	107	125
5	35	84	130	144

To illustrate, suppose Thomas is considering whether to purchase additional CNC machines. Engineers estimate the production function of additional machines and derive Table 5.3. The average product of either machine tools or machinists is computed by dividing the total output by the amount of either machine tools or machinists used. The marginal product of each input is obtained by holding the other input constant. For example, the marginal product of an additional machine tool when using four machinists and three machine tools is 5,100 parts per machine tool; the marginal product of an additional machinist when using three machinists and four CNC machines is 2,100 parts per unit. If X_1 is the amount of the first input and X_2 is the amount of the second input, the production function is

$$Q = f(X_1, X_2)$$

where Q is the firm's output rate. The marginal product of the first input is $\Delta Q / \Delta X_1$; the marginal product of the second input is $\Delta Q / \Delta X_2$.

QUANT OPTION

For the fastidious, we have

$$MP_1 = \frac{\partial Q}{\partial X_1} \quad MP_2 = \frac{\partial Q}{\partial X_2}$$

1. This surface is not meant to represent the numerical values in Table 5.3 but is a general representation of how a production surface of this sort is likely to appear.

Visually, we can represent the production function by a surface, as shown in Figure 5.3. The production surface is $OAQB$.¹ We measure output for any

STRATEGY SESSION: Substitution of Capital for Labor in Legal Proceedings

In lawsuits, the parties must submit documents to each other that are requested and ordered by the court. This enables each side to discover what the other side knows and to prepare an offense or defense against it. This process is called *discovery*. The traditional tactic used by the defense in large cases is to overload the plaintiff with documents, e-mails, and so on, including many materials the defense knows are irrelevant to the case. The plaintiff must then separate the wheat from the chaff—often a daunting task. Who performs the work? Often an army of paralegals and lawyers who are compensated handsomely (especially the lawyers) for many, many hours of work. In a classic example, one case (a U.S. Department of Justice antitrust case involving five television studios in 1978), 6 million documents were examined at a cost of \$2.2 million.

Enter the world of artificial intelligence (AI) and the computer. What's called "e-discovery" software can quickly scan documents and e-mails for key words and phrases. In addition, the software can point to patterns of behavior over multiple documents and develop timelines of who knew what when. In January of 2011, Blackstone Discovery helped analyze 1.5 million documents for a price that was less than \$100,000. In 2010, Clearwell (an e-discovery firm) analyzed 570,000 documents in two days and distilled them to 3,070 documents that were relevant to the case. The labor savings can be on the order of 99%; in other words, one lawyer can do what before had taken 100 to do. Mike Lynch, the founder of Autonomy (another e-discovery firm), predicts that one lawyer will soon be able to do the work of 500, and that new generations of software will enable one lawyer to do the work of 1,000.

Interestingly enough, good develops from bad. The rapid development of AI for the legal community came from the massive amount of documentation the U.S. Department of Justice accumulated in the Enron fraud case. More than 5 million pieces of

documentation were involved. When researchers got their hands on the documentation, they learned how participants conversed about nefarious activities and how social networks are formed.

Here's how sophisticated the programs are becoming. The programs focus on "linguistics" and "sociology." Linguistics has moved from just searching for keywords to linking those words to related words: synonyms, antonyms, activities associated with a word, and so on. An example might be that a search for cat would also give you *feline*, *meow*, *cuddly*, and *mice*. Sociology traces who's talking with whom about what, can trace changes in communications in terms of tone and content, and can even analyze whether certain words, phrases, or documents indicate stress levels of the sender or receiver. Examples here are when communications turn from casual to stiff or communication length/frequency changes. Words like *let's continue this offline* usually indicate that something is going on.

AI has penetrated into the computer industry itself. Software has replaced logic designers and draftsmen in the computer chip industry, and it now performs many of the tasks of loan officers in banks and of tax accountants (think TurboTax). This puts middle management jobs at risk and the trend could continue to upper management jobs.

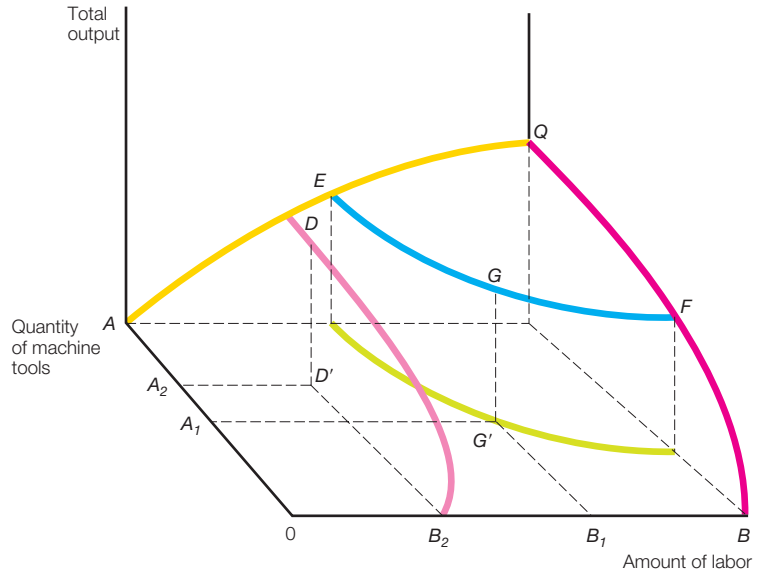
But one more question arises: The computer is fast, but is it good? A major chemical company lawyer tested e-discovery software on old cases of his company. He found that his company's lawyers were only 60% accurate relative to the computer. He points out that people get bored, get distracted, don't feel well, have bad days, and so on, but computers don't.

Source: John Markoff, "Armies of Expensive Lawyers, Replaced by Cheaper Software," *New York Times*, at www.nytimes.com/2011/03/05/science/05legal.

FIGURE 5.3

Production Function, Two Variable Inputs

The production surface, $OAQB$, shows the amount of total output that can be obtained from various combinations of machine tools and labor.



input bundle as height on the surface. Dropping a perpendicular down from a point on the surface to the “floor” defines the corresponding input bundle. For example, producing $G'G$ units of output requires $OB_1 (= A_1G')$ machinists and $OA_1 (= B_1G')$ machine tools. Conversely, we can take any amounts of machine tools and machinists, say OA_2 machine tools and OB_2 machinists, and find their output levels by measuring the height of the production surface at D' , the point where machinists is OB_2 and machine tool input is OA_2 . According to Figure 5.3, the output equals $D'D$. Input bundles that produce identical output have the same height.

ISOQUANTS

Isoquant Curve showing all possible (efficient) input bundles capable of producing a given output level.

An **isoquant** is a curve showing all possible (efficient) input bundles capable of producing a given output level. An isoquant is composed of all the points having the same height in the production surface of Figure 5.3. Suppose we want to find the isoquant corresponding to an output of $G'G$. All we need to do is cut the surface at the height of $G'G$ parallel to the base plane, the result being EGF , and drop perpendiculars from EGF to the base. Clearly this results in a curve that includes

STRATEGY SESSION: How Nucor Stays on the Production Function

Nucor is the largest steel firm in the United States, although it did not focus on steel production until the 1960s. Net sales for the first 39 weeks of 2011 were over \$15 billion. More remarkably, the company has paid a cash dividend for 154 consecutive quarters. This performance has far outstripped that of more traditional (integrated) steel manufacturers like U.S.X. For example, when the average integrated steel company in the United States produced 400 tons of steel per employee, Nucor produced about 980 tons per employee. What actions have Nucor's managers implemented to achieve this superior performance?

One difference is that Nucor is a "minimill," not an integrated steel firm. Minimills have a different production function than do integrated mills. They use electric arc furnaces to make steel products from scrap metal. In 2007 Nucor was the nation's largest recycler, reprocessing one ton of steel every two seconds.

The primary reason for Nucor's outstanding performance is the company's focus on the efficient use of resources. Management has executed strategies that keep the company on an efficient production function. Being efficient is not caused by random luck. It is clearly governed by managerial decisions and requires an integrated set of policies. In Nucor's case, the firm employs roughly 12,000 people. How do Nucor managers keep employees focused on efficient production?

Nucor uses the following multipronged approach:

1. It maintains a simple, streamlined organizational structure that encourages decentralized decision making. Most divisions use only three layers of management. Each division is treated as a profit center and is expected to earn a 25% return on total assets.
2. The company acts as the general contractor in building new plants. It locates plants in rural areas where land is cheap (and unions are weak). Also, each plant is located near water and is served by at least two railroad lines to keep freight rates low.

Nucor recruits employees to help build the plant; this allows it to observe the work habits of individuals. Those with good work habits are recruited to work in the plant when it opens. It also brings workers from other plants (who have already built plants) to join the construction team. Using these methods, Nucor can build a plant at a lower cost and in about 33% less time than competitors.

3. All employees are subject to performance-related compensation plans. For example, production employees are paid weekly bonuses based on the productivity of their work group. Using this team approach, Nucor can lower its monitoring costs because employees monitor each other. Bonuses are based on the capabilities of the equipment and average 80% to 150% of an employee's base pay. The more output a team produces, the higher are its bonuses.

4. Nucor treats all employees equally. Benefits are the same regardless of organizational position. There are no company cars, executive dining rooms, or corporate jets.

5. The firm's focus on output does not mean that quality suffers. Employees are committed to providing the highest-quality steel at a competitive price. To reinforce this commitment to quality, most of Nucor's divisions are ISO 9000 certified.^a

6. Finally, Nucor regards itself as a technological leader. It was the first firm to produce thin-slab casting at a minimill and searches worldwide for new developments in steel production. This emphasis on innovation is reinforced by the firm's flat organizational structure. Decisions can be made and implemented quickly.

^aISO 9000 is a set of quality standards. To receive ISO 9000 certification, managers must fulfill various quality assurance requirements and be audited by an external registrar. If a firm's quality assurance system is approved by this registrar, the firm is awarded an ISO 9000 certification and is allowed to advertise this fact to all customers.

all efficient combinations of machine tools and machinists that can produce G' metal parts. Using the notation in equation (5.1), an isoquant shows all combinations of X_1 and X_2 such that $f(X_1, X_2)$ equals a certain output.

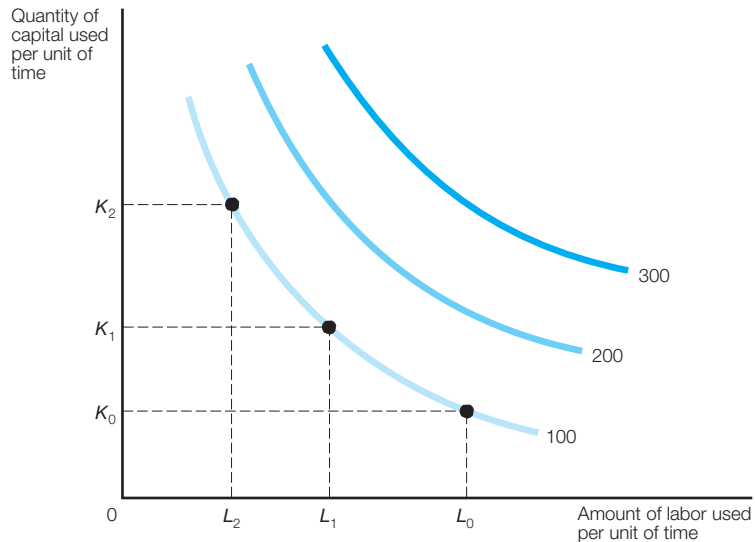
Several isoquants, each pertaining to a different output rate, are shown in Figure 5.4. The two axes measure the quantities of inputs. In contrast to the previous diagrams, we assume labor and capital—not machinists and machine tools (a particular form of labor and capital)—are the relevant inputs. The curves show the various input bundles that produce 100, 200, and 300 units of output. For example, consider the isoquant for 100 units of output. This isoquant shows it is possible to produce 100 units if L_0 units of labor and K_0 units of capital are used per time period. Alternatively, this output rate can be attained with L_1 units of labor and K_1 units of capital—or L_2 units of labor and K_2 units of capital.

Figure 5.4 illustrates several properties of isoquants. The farther the isoquant is from the origin, the greater the output it represents. Because we assume continuous production functions, we can draw an isoquant for any input bundle. Each isoquant represents an infinite number of possible input combinations. Isoquants are always downward-sloping and convex to the origin (we will see why in the next section).

FIGURE 5.4

Isoquants

These three isoquants show the various combinations of capital and labor that can produce 100, 200, and 300 units of output.



STRATEGY SESSION: Substitution

Firms now have manager job titles that weren't around a few years ago. One of those titles reads something like Chief Innovation Officer. A search of LinkedIn found at least 700 people who claimed this title in their firm and that an additional 25,000 had the word innovation somewhere in their job title. And who knows how many other managers innovate under a different job title. So what's new in innovation? The developments seem to fall into three areas: (1) Take a wide view and think out of the box; (2) Think entrepreneurially like those guys in the Silicon Valley; and (3) Listen to what customers are saying that they want and are missing and try to serve these needs. Thus, the innovators wish to explore new goods and services and how to produce them and existing goods and services more efficiently.

For instance, consider developments (1) and (2) from the previous paragraph. Weyerhaeuser (which produces paper and pulp) has discovered a new use for a chemical compound called *lignin*. It's a by-product of trees from pulp making and had been recycled as a fuel in the process of paper making. But some analysis showed that it could be used to produce the carbon fiber that automakers use to replace metal (to lower car weight and thus improve fuel efficiency). Lignin is worth 10 to 20 times more in the production of carbon fiber than it is as a fuel in a

paper plant. The automobile industry is changing its production function by substituting carbon fiber for metal, and the pulp industry is changing its production function by substituting another fuel for lignin.

Citigroup's Chief Innovation Officer is applying development (3) and changing the bank's focus to customers and away from banking products, asking what do customers want rather than asking what Citibank wants to sell? Affluent customers were grouped into four affinity groups ranging from up-and-comers (30 year olds) to retiring baby boomers and two groups in between. Then customer service and marketing were retooled to what those customers wanted.

The new banks are digital, featuring touchscreen workstations and videoconferencing links to financial experts. The new banks are virtually paperless (traditional banks use as many as 100 forms). Thus, capital has replaced labor, and virtual banking has replaced paper banking. The production function for banking is changing.

Source: Steve Lohr, "Innovate, Yes, but Make It Practical," *New York Times*, at www.nytimes.com/2010/08/15/business/15unboxed.html.

THE MARGINAL RATE OF TECHNICAL SUBSTITUTION

Generally a particular output can be produced with a number of input bundles. As we move along a particular isoquant, the **marginal rate of technical substitution (MRTS)** shows the rate at which one input is substituted for another (with output remaining constant). If, as in equation (5.1), the output produced is a function of two inputs,

$$Q = f(X_1, X_2)$$

Marginal rate of technical substitution (MRTS) MRTS shows the rate at which one input is substituted for another (with output remaining constant).

MRTS is

$$MRTS = -\frac{\Delta X_2}{\Delta X_1} \quad (5.2)$$

given that Q (output) is held constant.

Geometrically the marginal rate of technical substitution is -1 times the slope of the isoquant. This makes sense because $\Delta X_2/\Delta X_1$ measures the slope, which is downward or negative (so X_2 is on the y axis and X_1 is on the x axis).

It is useful for managers to think of MRTS as the ratio of marginal products, MP_1/MP_2 , for inputs 1 and 2. Managers need to be efficient. The marginal product metric shows the incremental effect on output of the last unit of input. In spirit, managers want to increase the use of inputs with relatively high marginal products, though they must also consider the costs of inputs.

The rate of substitutability between inputs is varied. In some production processes, one type of labor is easily substituted for another; in others, specialized inputs are required. In extreme cases, no substitution among inputs is possible; to produce a unit of output, a fixed amount of each input is required, and inputs must be used in fixed proportions. Figure 5.5 shows the firm's isoquants in such a case; as you can see, they are right angles. Few production processes allow no substitution among inputs, but in some, substitutability is limited. If perfect substitutability of inputs is possible, isoquants are straight lines connecting the two axes.

QUANT OPTION

It's time for some fun!

$$dQ = \left(\frac{\partial Q}{\partial X_1}\right)dX_1 + \left(\frac{\partial Q}{\partial X_2}\right)dX_2 = 0$$

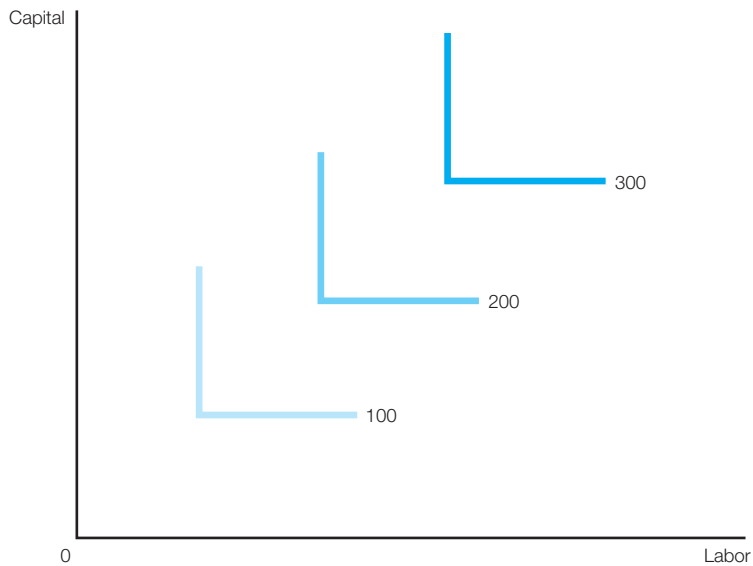
Therefore,

$$\frac{dX_2}{dX_1} = \frac{-(\partial Q/\partial X_1)}{\partial Q/\partial X_2} = -\frac{MP_1}{MP_2} \quad (5.3)$$

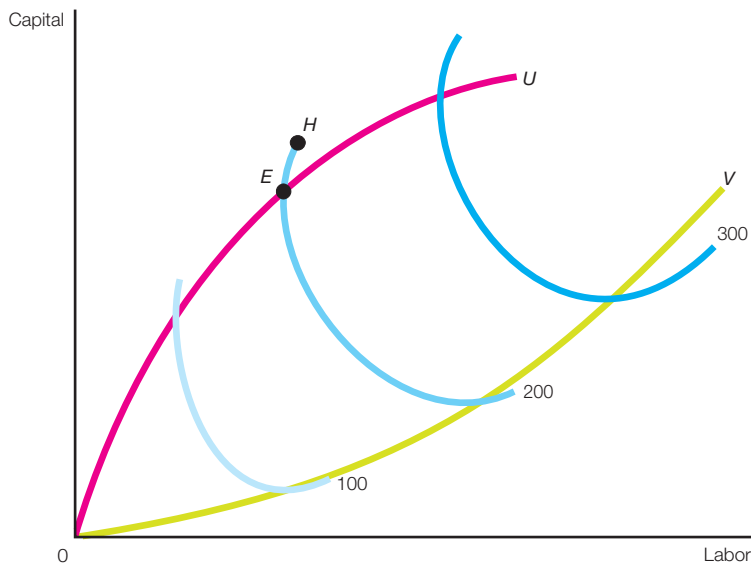
Mathematically, isoquants may have positively sloped segments or bend back on themselves, as shown in Figure 5.6. Above OU and below OV , the isoquant slopes are positive, implying that increases in both capital and labor are required to maintain a specified output rate. If this is the case, the marginal product of one or the other input is negative. Above OU , the marginal product of capital is negative; therefore, output increases if less capital is used while the level of labor is held constant. Below OV , the marginal product of labor is negative; output increases if

FIGURE 5.5**Isoquants in the Case of Fixed Proportions**

If inputs must be used in fixed proportions, the isoquants are right angles.

**FIGURE 5.6****Economic Region of Production**

No profit-maximizing firm operates at a point outside the ridge lines, OU and OV .



STRATEGY SESSION: Innovations in Payments for Goods and Services

Is the cash register a thing of the past? Already “cashiers” don’t key in item prices—UPC codes are scanned. Cashiers may take money and make change, but many payments are made by credit and debit cards. Some stores such as Home Depot, are already offering self-checkout lines. What’s next? How about a big arch near the exit that you walk your cart through that scans all your items purchased and puts the total on your credit card or reduces your bank balance by the amount of your purchases? It’s like an E-Z Pass for groceries or purchases from other retailers.

But how about paying by just swiping your smartphone? Deutsche Telekom, France Télécom, Google, Visa, and other companies are scrambling to get into what are called *near-field communication (NFC)* payment systems that will enable consumers to shop at the corner store and pay with a swipe of the phone. The telecoms are hoping that their existing payments system and billing arrangements with customers will be a first-mover advantage against

technology giants like Google (who bills virtually no one). The NFC market is estimated to account for up to a third of the global market in mobile transactions by 2014 (a \$1.3 trillion market). Not only cash registers but credit cards may be replaced.

The arrival of Google’s Android and Samsung’s Nexus S, as well as speculation that Apple (which knows how to bill) may enter the NFC market, make the future in this area very interesting.

Infrastructure investment will be significant. Gazillions of point of sales terminals must be put in place and then all phones will have to be NFC compatible. The most likely way all this will come about is through some partnership between the telecom experts and the financial services industry.

Source: Matthew Campbell and Jonathan Browning, “‘Mega’ Payments Race Pits Google, Visa Against Phone Operators,” *Bloomberg News*, February 21, 2011, at www.bloomberg.com/news/2011-02-20/-mega-payments-race-pits-google-visa-against-phone-operators.html.

Ridge lines The lines that profit-maximizing firms operate within, because outside of them, marginal products of inputs are negative.

less labor is used while the amount of capital is held constant. The lines OU and OV are called **ridge lines**.

No profit-maximizing manager will operate at a point outside the ridge lines because she can produce the same output with less of both inputs. This choice is strictly less costly. Consider point H in Figure 5.6. This point is located on a positively sloped segment of the isoquant (and so outside the ridge lines). It will always require greater levels of both labor and capital than a point inside the ridge lines (for example, point E) on the same isoquant. Because both capital and labor have positive prices, it is cheaper to operate at point E than at point H . The moral is this: Managers cannot use input bundles outside the ridge lines if they want to maximize profit.

THE OPTIMAL COMBINATION OF INPUTS

The previous analysis did not include the costs of inputs. Managers must consider costs because the inputs are scarce. A manager who wants to maximize profit will try to minimize the cost of producing a given output or maximize the output

derived from a given level of cost.² Suppose a manager takes input prices as given and uses two inputs, capital and labor, that vary in the relevant period. What combination of capital and labor should the manager choose to maximize the output derived from the given level of cost?

First we determine the various input combinations that can be obtained for a given cost. If capital and labor are the inputs and the price of labor is P_L per unit and the price of capital is P_K per unit, the input combinations that are obtained for a total outlay of M are such that

$$P_L L + P_K K = M \quad (5.4)$$

where L is the level of labor and K is the level of capital. Given M , P_L , and P_K , it follows that

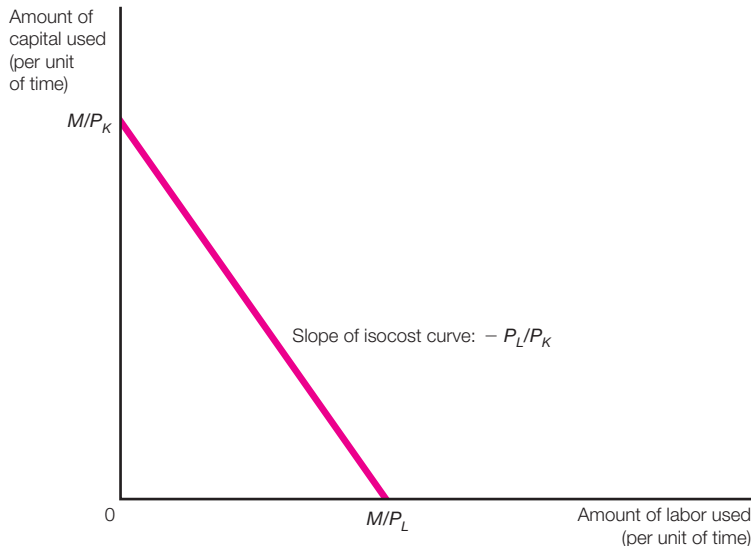
$$K = \frac{M}{P_K} - \frac{P_L L}{P_K} \quad (5.5)$$

The various bundles of capital and labor that can be purchased, given P_L , P_K , and M , are represented by the straight line shown in Figure 5.7. (Capital is plotted on the vertical axis, and labor is plotted on the horizontal axis.) This line, which has an intercept on the vertical axis equal to M/P_K and a slope of $-P_L/P_K$, is called an

FIGURE 5.7

Isocost Curve

The isocost curve shows the combinations of inputs that can be obtained for a total outlay of M .



2. The conditions for minimizing the cost of producing a given output are the same as those for maximizing the output from a given cost. This is shown in the present section. Therefore, we can view the firm's problem in either way.

Isocost curve Curve showing all the input bundles that can be purchased at a specified cost.

isocost curve. It shows all the input bundles that can be purchased at a specified cost (M).

If we superimpose the relevant isocost curve on the isoquant map, we see the input bundle that maximizes output for a given cost. An efficient manager should choose the point on the isocost curve that is tangent to the highest-valued isoquant—for example, R in Figure 5.8. Because the slope of the isocost curve is the negative of P_L/P_K and the slope of the isoquant is the negative of MP_L/MP_K (as we pointed out in the previous section), it follows that the optimal combination of inputs is one where $MP_L/MP_K = P_L/P_K$. Put differently, the firm should choose an input combination where $MP_L/P_L = MP_K/P_K$.

So efficient managers need to choose an input bundle where the marginal products per dollar spent of labor and capital are identical. If they are not, the manager should increase the use of the input with the higher marginal product per dollar value.

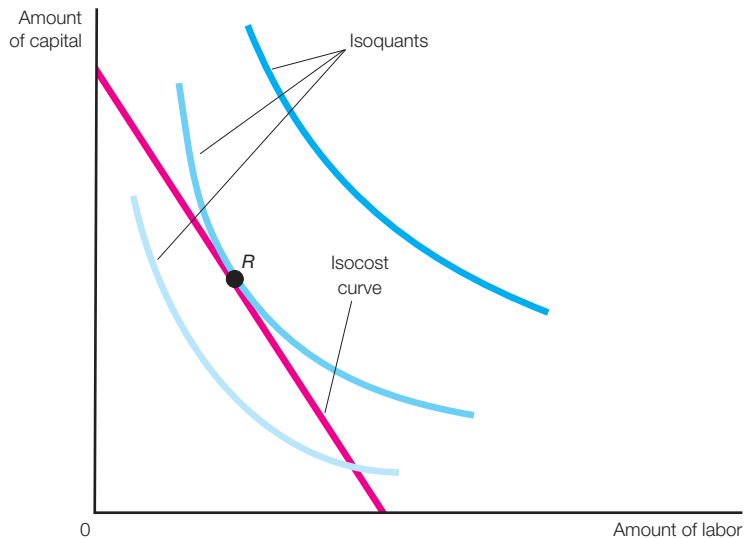
If there are more than two inputs, the manager maximizes output by distributing costs among the various inputs so the marginal product of a dollar's worth of one input is equal to the marginal product of a dollar's worth of any other input used. In spirit, the manager chooses an input bundle such that

$$\frac{MP_a}{P_a} = \frac{MP_b}{P_b} = \dots = \frac{MP_n}{P_n} \quad (5.6)$$

FIGURE 5.8

Maximization of Output for a Given Cost

To maximize the output for a given cost, the firm should choose the input combination at point R .



where MP_a, MP_b, \dots, MP_n are the marginal products of inputs a, b, \dots, n ; and P_a, P_b, \dots, P_n are the prices of inputs a, b, \dots, n .

To determine the input bundle that minimizes production costs, we use a graph similar to Figure 5.8. Moving along the isoquant of the stipulated output level, we find the point that lies on the lowest isocost curve—for example, S in Figure 5.9. Input bundles on isocost curves like C_0 that lie below S are cheaper than S , but they cannot produce the desired output. Input bundles on isocost curves like C_2 that lie above S produce the desired output but at a higher cost than S . It is obvious that the optimal bundle S is a point where the isocost curve is tangent to the isoquant. Therefore, to minimize the cost of producing a given output or to maximize the output from a given cost outlay, the firm must equate

STRATEGY SESSION: The Substitution of Inputs in Baseball

The regular baseball season is a grueling 162-game odyssey that leads eight teams to the playoffs and, ultimately, two teams to the World Series. Surviving the 162 games and arriving at the playoffs entails luck and skill. Pitching is a large percentage of the game. Most teams operate with a four or five starting-pitcher rotation (meaning that each pitcher starts between 30 and 40 games and gets about four days rest between starts). Over the long haul of the season, a pitcher needs that time between starts to avoid wear and tear on his arm.

But when teams make it to the playoffs, things change. Because of strategic scheduling of games and because there is only a short time until the pitchers can rest during the off-season, teams are able to go with their best three (or four) pitchers instead of using their fourth or fifth starters. By using their best pitchers, post season ERAs (earned run averages) have been lower in the post season than in the regular season in 13 of the last 15 years. In 2001, the ERA in the post season was 1.3 runs less than in the regular season (almost a 30% decrease).

Which pitchers are the best pitchers for the playoffs? The power pitchers, such as Josh Beckett, John Smoltz, “CC” Sabathia, Curt Schilling, and Randy Johnson are best. While some finesse pitchers

have done well, including Cliff Lee and Cole Hamels, other regular season finesse pitchers, such as Greg Maddux and Tom Glavine, have been subpar in the post season. The power pitchers tend to dominate during the playoffs.

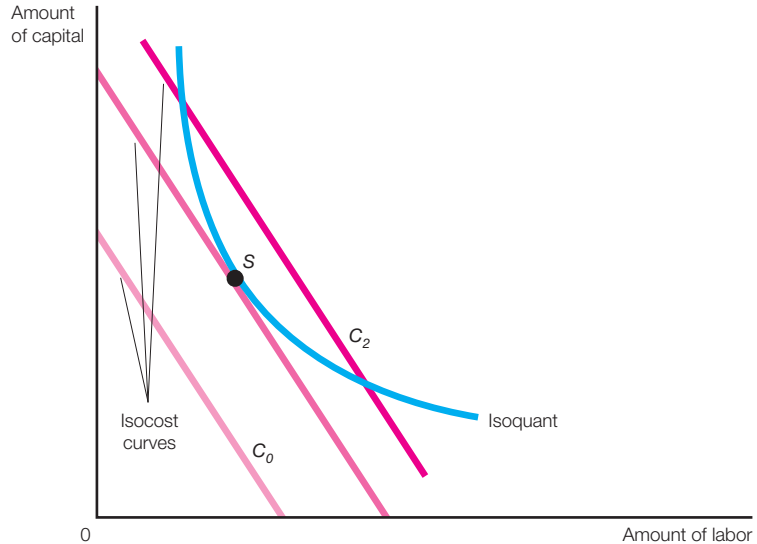
Why does management (i.e., the managers and coaches) substitute power for finesse? John Smoltz—a former power pitcher, playoff Most Valuable Player, and holder of a 15-4 playoff win-loss record and a 2.67 playoff ERA, and now a baseball color commentator—states, “If I had to orchestrate a team, the front end of my rotation would have power guys. Whether or not they were super successful during the regular season, they have a better chance of dominating in a short series.” Tampa Bay Ray’s pitching coach Jim Hickey says, “I think every team would prefer to go with power arms. If you get a dominant pitcher who strikes people out, there are no other variables.” Power pitchers control their destiny with strikeouts whereas finesse pitchers let batters make contact with the ball and hence create the possibility of hits and errors.

Source: Jorge L. Ortiz, “Pitching Rules in Playoffs: ERAs Drop, Strikeouts Go Up,” *USA Today*, October 5, 2010, at www.usatoday.com/sports/baseball/playoffs/2010-10-05.

FIGURE 5.9

Minimization of Cost for a Given Output

To minimize the cost of producing the amount of output corresponding to this isoquant, the manager should choose the input combination at point *S*.



MP_L/MP_K and P_L/P_K ; this means that $MP_L/P_L = MP_K/P_K$. And if more than two inputs are needed, the manager must satisfy equation (5.6).

CORNER SOLUTIONS

Just as in consumption theory in Chapter 3, it is possible to have corner solutions—that is, optimal input bundles with just one input deployed. Given the production technology and the prices of the inputs, there may be no tangency of an isoquant with an isocost curve. In the two-input case, this means that just one input is used to produce the product in the least expensive way (or to produce the most output with a given cost). Equation 5.6 will now be an inequality reading $MP_K/P_K > MP_L/P_L$ for cases where just capital is used and $MP_K/P_K < MP_L/P_L$ for cases where just labor is used. The former case is shown in Figure 5.10.

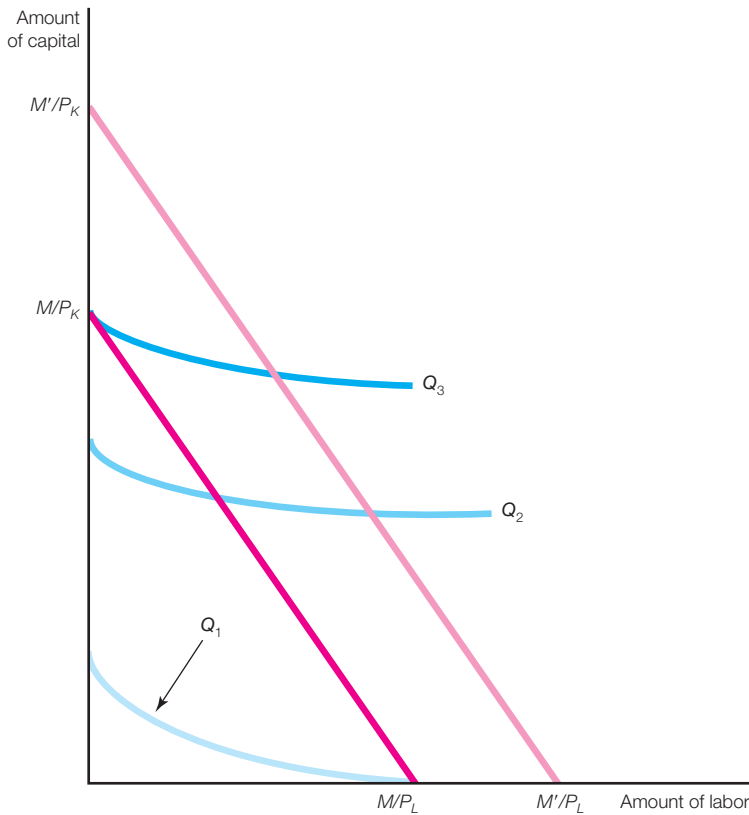
RETURNS TO SCALE

We have seen how managers can represent technology as a production function and use concepts like marginal and average product to operate more efficiently. We want to continue this theme and examine some long-term considerations

FIGURE 5.10

A Corner Solution Where Only One Input Is Used

With outlay of M , the most that can be produced is Q_3 using only capital (M/P_K units). If only labor were used, the firm could produce only Q_1 units with outlay M . The cheapest way to produce Q_3 units is with just M/P_K units of capital and with no labor. Q_3 units could be produced with an outlay of $M' > M$ by using both capital and labor, but that would be inefficient.



managers face. These focus on scale. Basically, what is the incremental change to output as managers increase their use of capital and labor?

Suppose we consider a long-term situation where all inputs are variable, and managers increase the level of inputs by the same proportion. What will happen to output? Clearly there are three possibilities. First, output may increase by a larger proportion than inputs; for example, doubling all inputs may more than double output. This is the case of **increasing returns to scale**. Or output may increase by a smaller proportion than inputs; for example, doubling all inputs

Increasing return to scale When output increases by a larger proportion than inputs.

Decreasing returns to scale

When output increases by a smaller proportion than inputs.

Constant returns to scale

When output increases by exactly the same proportion as inputs.

may lead to less than a doubling of output. This is the case of **decreasing returns to scale**. Finally, output may increase by exactly the same proportion as inputs; for example, doubling all inputs may double output. This is the case of **constant returns to scale**.

At first glance, some managers may believe that production functions necessarily exhibit constant returns to scale. After all, if a manager can build two factories with the same plant size and types of workers, can't she achieve the same output with a single plant twice the size? But things are not this simple. If managers double the size of a plant, they may employ techniques that are economically infeasible at the smaller scale. Some inputs are not available in small units; for example, we

PROBLEM SOLVED: What Skills Do We Need?

Consider the Beiswanger Company, a small firm engaged in engineering analysis. Beiswanger's president has estimated that the firm's output per month (Q) is related in the following way to the number of engineers (E) and technicians used (T)

$$Q = 20E - E^2 + 12T - 0.5T^2 \quad (5.7)$$

The monthly wage of an engineer is \$4,000, and the monthly wage of a technician is \$2,000. If the president allots \$28,000 per month for the combined wages of engineers and technicians, what mix of engineers and technicians should he hire?

If the president is to maximize output (for his \$28,000 budget), he must choose a bundle of engineers and technicians such that

$$\frac{MP_E}{P_E} = \frac{MP_T}{P_T} \quad (5.8)$$

where MP_E is the marginal product of an engineer, MP_T is the marginal product of a technician, P_E is the wage of an engineer, and P_T is the wage of a technician. Viewing the change in equation (5.7) with respect to E and T , we find that

$$MP_E = \frac{\Delta Q}{\Delta E} = 20 - 2E \quad (5.9a)$$

$$MP_T = \frac{\Delta Q}{\Delta T} = 12 - T \quad (5.9b)$$

Inserting these expressions for MP_E and MP_T into equation (5.8) and noting that $P_E = 4,000$ and $P_T = 2,000$, it follows that

$$\begin{aligned} \frac{20 - 2E}{4,000} &= \frac{12 - T}{2,000} \\ \frac{2,000(20 - 2E)}{4,000} &= 12 - T \\ 10 - E &= 12 - T \\ T &= E + 2 \end{aligned}$$

Because Beiswanger allocates \$28,000 per month for the total wages of engineers and technicians, we have

$$4,000E + 2,000T = 28,000$$

Substituting $(E + 2)$ for T gives us

$$4,000E + 2,000(E + 2) = 28,000$$

This means that $E = 4$ (and $T = 6$). So to maximize output from the \$28,000 outlay on wages, the president should hire four engineers and six technicians.

PROBLEM SOLVED: The Efficient Minds of Managers

Managers need to search for input bundles to minimize costs for a given output. Intuitively they need to balance the productivity of an input with its cost. As we will show in Chapter 6, there are easier metrics for managers to use, like costs. But the ability to control costs is enhanced by understanding production functions. Consider the issues facing managers at the Miller Company, for which the relationship between output per hour (Q) and the number of workers (L) and machines (K) used per hour is

$$Q = 10(LK)^{0.5}$$

The wage of a worker is \$80 per hour, and the price of a machine is \$20 per hour. If the Miller Company produces 800 units of output per hour, how many workers and machines should managers use?

According to equation (5.8), the Miller Company should choose an input combination such that

$$\frac{MP_L}{P_L} = \frac{MP_K}{P_K}$$

where MP_L is the marginal product of a worker, MP_K is the marginal product of a machine, P_L is the wage

of a worker, and P_K is the price of using a machine. Because $Q = 10(LK)^{0.5}$,

$$MP_L = \frac{\Delta Q}{\Delta L} = 5\left(\frac{K}{L}\right)^{0.5}$$

$$MP_K = \frac{\Delta Q}{\Delta K} = 5\left(\frac{L}{K}\right)^{0.5}$$

So if $MP_L/P_L = MP_K/P_K$

$$\frac{5(K/L)^{0.5}}{80} = \frac{5(L/K)^{0.5}}{20}$$

Multiplying both sides of this equation by $(K/L)^{0.5}$, we get

$$\frac{5K}{80L} = \frac{5}{20}$$

which means that $K = 4L$. Because $Q = 800$,

$$10(LK)^{0.5} = 800$$

$$10[L(4L)]^{0.5} = 800$$

$$L = 40$$

$$K = 160$$

Therefore, to minimize cost, managers at the Miller Company should hire 40 workers and use 160 machines.

cannot install half a robot. Because of indivisibilities of this sort, larger plants may have increasing returns to scale. So when managers think about efficient choice, building one large factory relative to two smaller ones may be better.

Larger plants also let managers subdivide tasks and use inputs more narrowly. This specialization strategy increases production efficiency; so investment bankers specialize in designated areas, and airlines dedicate computers to handling reservations. Larger size may also generate probabilistic efficiencies; for example, because the aggregate behavior of a larger number of customers tends to be more stable, a firm's inventory may not have to increase in proportion to its sales.

Increasing returns to scale also arise because of certain geometrical relations. Because the volume of a box that is $2 \times 2 \times 2$ feet is eight times as great as the volume of a box that is $1 \times 1 \times 1$ foot, the former box can carry eight times as much as the latter box. But the area of the six sides of the $2 \times 2 \times 2$ box is 24 square feet and the area of the six sides of the $1 \times 1 \times 1$ box is six square feet, so the former box requires only four times as much wood as the latter.

It turns out that bigger is not always better; managers can experience decreasing returns to scale. The most common culprit is the challenge of coordinating a large organization. It can be difficult even in a small firm for managers to obtain the necessary information to make important decisions; in a large firm such problems tend to be greater. As we show in Chapter 15, managers often have difficulties in designing efficient incentive schemes in larger firms. Though the advantages of a large organization are obvious, scale can generate inefficiencies. For example, in certain kinds of research and development, large teams tend to be less effective than smaller ones.

Whether scale returns are constant, increasing, or decreasing is an empirical question that must be settled case by case. There is no simple, all-encompassing answer. In some industries the evidence suggests that returns increase over a certain range of output; but the answer is likely to depend on the output considered. There may be increasing returns to scale at small output levels and constant or decreasing returns to scale at higher levels. In addition, managers need to know how output changes when inputs are not all increased or decreased in the same proportion.

THE OUTPUT ELASTICITY

To measure whether there are increasing, decreasing, or constant returns to scale, the output elasticity is computed. The **output elasticity** is defined as the percentage of change in output resulting from a 1% increase in all inputs. If the output elasticity exceeds 1, there are increasing returns to scale; if it equals 1, there are constant returns to scale; and if it is less than 1, there are decreasing returns to scale.

As an illustration, consider the Lone Star Company, a maker of aircraft parts, which has the following production function

$$Q = 0.8L^{0.3}K^{0.8}$$

Here Q is the number of parts produced per year (measured in millions of parts), L is the number of workers hired, and K is the amount of capital used. This is the commonly used Cobb-Douglas production function (named after Charles Cobb and Paul Douglas, who pioneered its application).

Output elasticity The percentage of change in output resulting from a 1% increase in all inputs.

To calculate the output elasticity at the Lone Star Company, let's see what will happen to Q if we multiply both inputs (L and K) by 1.01. Clearly the new value of Q (that is, Q') equals

$$\begin{aligned} Q' &= 0.8(1.01L)^{0.3}(1.01K)^{0.8} \\ &= 0.8(1.01)^{1.1}L^{0.3}K^{0.8} \\ &= (1.01)^{1.1}(0.8L^{0.3}K^{0.8}) \\ &= (1.01)^{1.1}Q \\ &= 1.011005484Q \end{aligned}$$

Therefore, if a manager increases the use of both inputs by 1%, output increases by slightly more than 1.1%; this means the output elasticity is approximately 1.1. It is exactly 1.1 for an infinitesimal change in input use (of both inputs). Because a 1% change is larger than infinitesimal, the increase in output is slightly larger than 1.1.

ESTIMATIONS OF PRODUCTION FUNCTIONS

Managers need to estimate production functions. One of the first steps in estimating a production function is to choose its mathematical form. Managers commonly use the Cobb-Douglas form. With only two inputs, this form is

$$Q = aL^bK^c \quad (5.10)$$

where Q is the number of parts produced per year (measured in millions of parts), L is the number of workers hired, and K is the amount of capital used. One advantage of this form is that the marginal productivity of each input depends on the level of all inputs employed, which is often realistic. Consider the marginal product of labor, which equals

$$\frac{\Delta Q}{\Delta L} = baL^{b-1}K^c = b\left(\frac{Q}{L}\right) = b(AP_L)$$

QUANT OPTION

$$\partial Q / \partial L = baL^{b-1}K^c = baL^{b-1}K^c(L/L) = baL^bK^c/L = b(Q/L) = b(AP_L)$$

Obviously the marginal product of labor depends on the values of both L and K . Another advantage is that if logarithms are taken of both sides of equation (5.10),

$$\log Q = \log a + b \log L + c \log K \quad (5.11)$$

PROBLEM SOLVED: Finding the Optimal Mix

Consider the production of broiler chickens, which is a big industry in the United States (2010 production value: \$20.8 billion). At one company, managers ran an experiment in which broilers were fed various amounts of corn and soybean oilmeal and the gain in weight of each broiler was measured. The managers then used regression to estimate the production function for broilers

$$G = 0.03 + 0.48C + 0.64S - 0.02C^2 - 0.05S^2 - 0.02CS \quad (5.12)$$

Here G is the gain in weight (in pounds per broiler), C is pounds of corn per broiler, and S is pounds of soybean oilmeal per broiler. The multiple coefficient of determination (R^2) is very high—about 0.998.

Using equation (5.12), managers can estimate isoquants for poultry production. Suppose they want to estimate the isoquant pertaining to a weight gain of one pound. In other words, they want to find the various combinations of corn per broiler and soybean oilmeal per broiler that results in a weight gain per broiler of one pound. To find these combinations, set $G = 1$

$$1 = 0.03 + 0.48C + 0.64S - 0.02C^2 - 0.05S^2 - 0.02CS \quad (5.13)$$

Then we set C equal to various values and determine each resulting value of S . For example, suppose $C = 1$. Then

$$1 = 0.03 + 0.48(1) + 0.64S - 0.02(1^2) - 0.05S^2 - 0.02(1)S$$

or

$$1 = 0.03 + 0.48 - 0.02 + (0.64 - 0.02)S - 0.05S^2$$

Solving $0.05S^2 - 0.62S + 0.51 = 0$ by the quadratic formula yields

$$\begin{aligned} S &= [0.62 \pm \sqrt{(0.62^2 - 4(0.05)(0.51))}] / 2(0.05) \\ &= [0.62 \pm \sqrt{(0.3844 - 0.102)}] / 0.1 \\ &= [0.62 \pm \sqrt{0.2824}] / 0.1 \\ &= [0.62 \pm 0.5314] / 0.1 \end{aligned}$$

Therefore, $S = 1.1514 / 0.1 = 11.514$, or $S = 0.08858 / 0.1 = 0.886$. Consequently, if a broiler is to gain one pound of weight, it must be fed 0.886 pounds of soybean oilmeal, as well as one pound of corn.^a

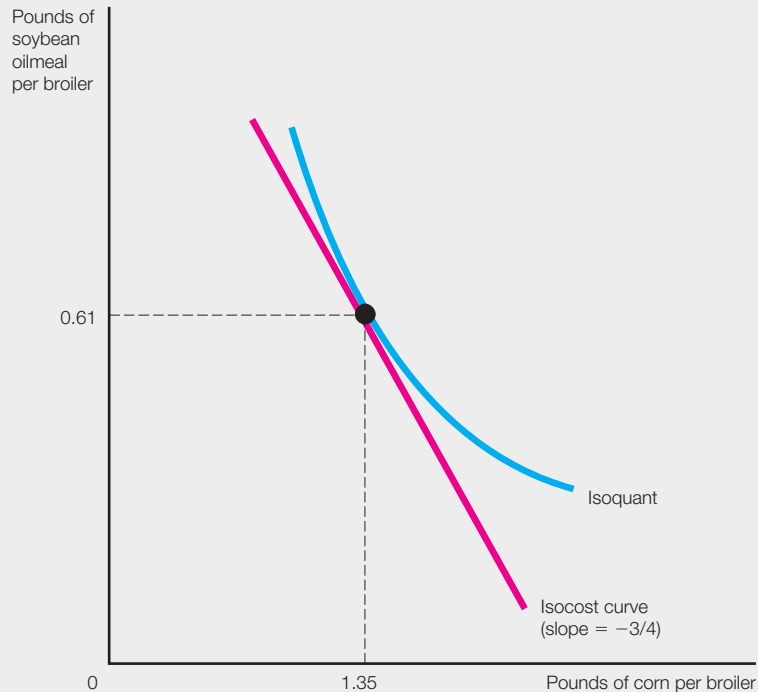
If we let $C = 1.1$, we can find the corresponding value of S by substituting 1.1 for C in equation (5.13) and solving for S . If we let $C = 1.2$, we can find the corresponding value of S by substituting 1.2 for C in equation (5.13) and solving for S . Proceeding in this way, we can find more and more points on the isoquant corresponding to a weight gain of one pound. The resulting isoquant is shown in the figure. Isoquants of this sort are of great importance to managers. Coupled with data regarding input prices, they can be used to determine which input bundles will minimize costs (recall Figure 5.9).

Managers use the isoquant in the figure shown on page 163 to determine how much corn and soybean oilmeal to feed a broiler if they want a one-pound weight gain. To see how, suppose the price of a pound of corn is three-quarters the price of a pound of soybean oilmeal. Then the slope of each isocost curve in the figure equals $-3/4$ because, as pointed out in Figure 5.7, the slope equals -1 times the price of the input on the horizontal axis (corn) divided by the price of the input on the vertical axis (soybean oilmeal). For the cost of the weight gain to be at a minimum, the isocost curve should be tangent to the isoquant; this means that the slope of the isoquant should also equal $-3/4$. As shown in the figure,

Isoquant for a One-Pound Weight Gain for a Broiler and Isocost Curve If Corn Price Is $\frac{3}{4}$ of Soybean Oilmeal Price

The optimal input combination is 1.35 pounds of corn and 0.61 pounds of soybean oilmeal.

Source: Organization for Economic Cooperation and Development, Interdisciplinary Research



this occurs when 1.35 pounds of corn and 0.61 pounds of soybean oilmeal are used. Therefore, this is the optimal input combination if the price of a pound of corn is three-quarters the price of a pound of soybean oilmeal.

^a There is another possible value of S , which corresponds to the use of the plus sign (rather than the minus sign) before $(b^2 - 4ac)^{0.5}$ in this formula; but this other value is not relevant here.

STRATEGY SESSION: Economies in Oil Tankers

Over 65% of crude oil output is transported by oil tankers. Oil is the largest commodity in transocean trade, accounting for 40% of all ocean shipments by weight. Oil tankers can be regarded as large cylinders. The surface area of a cylinder is not proportional to its volume; instead, as a cylinder's volume increases, its surface area goes up less than proportionately. Therefore, a tanker that can carry 300,000 deadweight tonnes (dwt) is only about twice as broad, long, and deep as one that can carry 30,000 tonnes.

Since the 1970s, the size of oil tankers has increased, as shown here:

Year	Average Oil Tanker Size (Thousands of dwt)
1973	64.0
1978	103.0
1985	146.0
2000	220.0
2008	273.0
2011	300.0

There is a strong cost incentive for larger tankers to be built. The cost of constructing an oil tanker is largely based on the cost of steel. A manager can increase the capacity of a tanker eight times by only using four times the amount of steel. A 280,000 dwt tanker costs roughly \$85 million to build, whereas a 28,000 dwt tanker costs roughly \$20 million to build.

Larger tankers can also operate with relatively smaller crews. An oil tanker of over 200,000 dwt today operates with a crew of 24—roughly half the crew size of a ship half that size in the 1980s. Onshore personnel costs are also relatively lower with larger tankers because these are based on the number of ships and not total tonnage. Finally, fuel costs of larger tankers are relatively lower. For example, a 60,000 dwt tanker generally requires 16,000 horsepower to travel at 15 knots. A 260,000 dwt tanker requires 42,500 horsepower to do the same. So 2.7 times the energy enables 4.3 times as much cargo to be shipped.

Sources: www.oceanatlas.com/unatlas/uses/transportation

Note that if managers use the Cobb-Douglas form, they can easily estimate the returns to scale. If the sum of the exponents (that is, $b + c$) exceeds 1, increasing returns to scale are indicated; if the sum of the exponents equals 1, constant returns to scale prevail; and if the sum of the exponents is less than 1, decreasing returns to scale are indicated. This is true because if the Cobb-Douglas production function prevails, the output elasticity equals the sum of the exponents. For example, in the previous section the output elasticity of the Lone Star Company was 1.1, which equaled the sum of the exponents (0.3 and 0.8).

There is no cut-and-dried way to determine which mathematical form is best because the answer depends on the particular situation. Frequently a good procedure is to try more than one mathematical form and see which fits the data best. The important thing is that the chosen form provide a faithful representation of

the actual situation. To determine whether this is the case, it often is useful to see how well a particular estimated production function can forecast the quantity of output resulting from the combination of inputs actually used.

SUMMARY

1. The production function defines the relationship among various inputs and the maximum quantity of a good that can be produced. Managers study production functions to gain insights into the firm's cost structure.
2. An isoquant is a curve showing all possible (efficient) combinations of inputs capable of producing a particular quantity of output. The marginal rate of technical substitution shows the rate at which one input can be substituted for another input if output remains constant. No profit-maximizing manager will operate at a point where the isoquant is positively sloped.
3. To minimize the cost of producing a particular output, a manager should allocate expenditures among various inputs so that the ratio of the marginal product to the input price is the same for all inputs used. Graphically, this amounts to choosing the input combination where the relevant isoquant is tangent to an isocost curve.
4. If a manager increases all inputs by the same proportion and output increases by more (less) than this proportion, there are increasing (decreasing) returns to scale. Increasing returns to scale may occur because of indivisibility of inputs, various geometrical relations, or specialization. Decreasing returns to scale can also occur; the most frequently cited reason is the difficulty of managing a huge enterprise. Whether there are constant, increasing, or decreasing returns to scale is an empirical question that must be settled case by case.
5. Managers have estimated production functions in many firms and industries. Many studies show that a Cobb-Douglas function is the best fit for the data.

PROBLEMS

1. In the Elwyn Company, the relationship between output (Q) and the number of hours of skilled labor (S) and unskilled labor (U) is

$$Q = 300S + 200U - 0.2S^2 - 0.3U^2$$

The hourly wage of skilled labor is \$10, and the hourly wage of unskilled labor is \$5. The firm can hire as much labor as it wants at these wage rates.

- a. Elwyn's chief engineer recommends that the firm hire 400 hours of skilled labor and 100 hours of unskilled labor. Evaluate this recommendation.
- b. If the Elwyn Company decides to spend a total of \$5,000 on skilled and unskilled labor, how many hours of each type of labor should it hire?

- c. If the price of a unit of output is \$10 (and does not vary with output level), how many hours of unskilled labor should the company hire?
2. A consulting firm specializing in agriculture determines that the following combinations of hay and grain consumption per lamb will result in a 25-pound gain for a lamb:

Pounds of Hay	Pounds of Grain
40	130.9
50	125.1
60	120.1
70	115.7
80	111.8
90	108.3
110	102.3
130	97.4
150	93.8

- a. The firm's president wants to estimate the marginal product of a pound of grain in producing lamb. Can he do so on the basis of these data?
- b. The firm's president is convinced that constant returns to scale prevail in lamb production. If this is true and hay and grain consumption per lamb are the only inputs, how much gain accrues if the hay consumption per lamb is 100 pounds and the grain consumption per lamb is 250.2 pounds?
- c. What is the marginal rate of technical substitution of hay for grain when between 40 and 50 pounds of hay (and between 130.9 and 125.1 pounds of grain) are consumed per lamb?
- d. A major advance in technology occurs that allows farmers to produce a 25-pound gain per lamb with less hay and grain than the preceding table indicates. If the marginal rate of technical substitution (at each rate of consumption of each input) is the same after the technological advance as before, can you draw the new isoquant corresponding to a 25-pound gain per lamb?
3. The Ascot Corporation, which produces stationery, hires a consultant to estimate its production function. The consultant concludes that

$$Q = 0.9P + 0.06L$$

where Q is the number of pounds of stationery produced by Ascot per year, L is the number of hours of labor per year, and P is the number of pounds of paper used per year.

- a. Does this production function seem to include all the relevant inputs? Explain.

- b. Does this production function seem reasonable if it is applied to all possible values of L ? Explain.
- c. Does this production function exhibit diminishing marginal returns?
4. A Cobb-Douglas production function was estimated for six types of farms. There were five inputs in the production function: (1) land, (2) labor, (3) equipment, (4) livestock and feed, and (5) other resource services. The exponent of each input was as follows:

Farm Type	Exponent				
	Land	Labor	Equipment	Livestock and Feed	Other Resource Services
Crop farms	0.24	0.07	0.08	0.53	0.02
Hog farms	0.07	0.02	0.10	0.74	0.03
Dairy farms	0.10	0.01	0.06	0.63	0.02
General farms	0.17	0.12	0.16	0.46	0.03
Large farms	0.28	0.01	0.11	0.53	0.03
Small farms	0.21	0.05	0.08	0.43	0.03

- a. Do there appear to be increasing returns to scale in any of these six types of farms?
- b. In what type of farm does a 1% increase in labor have the largest percentage effect on output?
- c. Based on these results, would you expect output to increase if many of the farms included in this sample were merged?
5. According to the chief engineer at the Zodiac Company, $Q = AL^\alpha K^\beta$, where Q is the output rate, L is the rate of labor input, and K is the rate of capital input. Statistical analysis indicates that $\alpha = 0.8$ and $\beta = 0.3$. The firm's owner claims the plant has increasing returns to scale.
- a. Is the owner correct?
- b. If β were 0.2 rather than 0.3, would she be correct?
- c. Does output per unit of labor depend only on α and β ? Why or why not?
6. According to data obtained by the U.S. Department of Agriculture, the relationship between a cow's total output of milk and the amount of grain it is fed is as follows:

Amount of Grain (Pounds)	Amount of Milk (Pounds)
1,200	5,917
1,800	7,250
2,400	8,379
3,000	9,371

- (This relationship assumes that forage input is fixed at 6,500 pounds of hay.)
- a. Calculate the average product of grain when each amount is used.
 - b. Estimate the marginal product of grain when between 1,200 and 1,800 pounds are fed, when between 1,800 and 2,400 pounds are fed, and when between 2,400 and 3,000 pounds are fed.
 - c. Does this production function exhibit diminishing marginal returns?
7. An electronics plant's production function is $Q = 5LK$, where Q is its output rate, L is the amount of labor it uses per period, and K is the amount of capital it uses per period. The price of labor is \$1 per unit of labor, and the price of capital is \$2 per unit of capital. The firm's vice president for manufacturing hires you to determine which combination of inputs the plant should use to produce 20 units of output per period.
 - a. What advice would you give him?
 - b. Suppose the price of labor increases to \$2 per unit. What effect will this have on output per unit of labor?
 - c. Is this plant subject to decreasing returns to scale? Why or why not?
 8. Volvo A.B., the Swedish auto firm, operated a car assembly plant at Uddevalla in 1988. The idea was to have a small team of highly skilled workers build an entire car. According to the proponents, this would reduce the tedium associated with the conventional assembly line and cut absenteeism and turnover among workers. In 1991 there were reports that it took 50 hours of labor to assemble a car at Uddevalla, in contrast to 25 hours at Volvo's conventional assembly plant at Ghent, Belgium. If you were Volvo's chief executive officer, what questions would you ask Uddevalla's managers, and what steps would you take?

APPENDIX: LAGRANGIAN MULTIPLIERS AND OPTIMAL INPUT COMBINATIONS

In this chapter we stated that equation (5.6) must be satisfied if a firm is to maximize output for a given expenditure level or if it is to minimize the cost of producing a specified amount of output. In this appendix we show how the decision rule in equation (5.6) is derived using the method of Lagrangian multipliers. To keep things relatively simple, we assume the manager is using only two inputs.

Maximizing Output from a Specified Expenditure Level

Suppose a firm's production function is

$$Q = f(X_1, X_2)$$

where Q is output, X_1 is the amount used of the first input, and X_2 is the amount used of the second input. The firm's total expenditure on both inputs is specified to equal E^* . Therefore,

$$X_1P_1 + X_2P_2 = E^*$$

where P_1 is the price of the first input and P_2 is the price of the second input. The manager seeks to maximize output for this specified level of expenditure. So she wants to maximize Q , where

$$Q = f(X_1, X_2) \quad (5.14)$$

subject to the constraint that

$$E^* - X_1P_1 - X_2P_2 = 0 \quad (5.15)$$

We first construct the Lagrangian function, which is the right side of equation (5.14) plus λ times the left side of equation (5.15)

$$L_1 = f(X_1, X_2) + \lambda(E^* - X_1P_1 - X_2P_2)$$

where λ is the Lagrangian multiplier. Taking the partial derivatives of L_1 with respect to X_1 , X_2 , and λ and setting them all equal to zero, we obtain

$$\frac{\partial L_1}{\partial X_1} = \frac{\partial f(X_1, X_2)}{\partial X_1} - \lambda P_1 = 0 \quad (5.16)$$

$$\frac{\partial L_1}{\partial X_2} = \frac{\partial f(X_1, X_2)}{\partial X_2} - \lambda P_2 = 0 \quad (5.17)$$

$$\frac{\partial L_1}{\partial \lambda} = E^* - X_1P_1 - X_2P_2 = 0 \quad (5.18)$$

These are the conditions for output maximization subject to the expenditure constraint.

MP_1 is the marginal product of input one and MP_2 is that for input two. By definition, we know the following is true

$$\frac{\partial f(X_1, X_2)}{\partial X_1} = \frac{\partial Q}{\partial X_1} = MP_1$$

$$\frac{\partial f(X_1, X_2)}{\partial X_2} = \frac{\partial Q}{\partial X_2} = MP_2$$

Equations (5.16) and (5.17) can be restated as

$$MP_1 - \lambda P_1 = 0$$

$$MP_2 - \lambda P_2 = 0$$

which implies that

$$MP_1 = \lambda P_1 \quad (5.19)$$

$$MP_2 = \lambda P_2 \quad (5.20)$$

Dividing each side of equation (5.19) by the corresponding side of equation (5.20), we find that

$$\frac{MP_1}{MP_2} = \frac{P_1}{P_2}$$

or

$$\frac{MP_1}{P_1} = \frac{MP_2}{P_2} \quad (5.21)$$

which is the decision rule in equation (5.6) when there are only two inputs. Thus we have shown why managers want to equate the marginal product per dollar spent across all inputs using the method of Lagrangian multipliers when the object is to maximize output subject to an expenditure constraint.

Minimizing the Cost of a Specified Amount of Output

Suppose a manager is committed to produce a specified quantity of output, Q^* , which means that

$$f(X_1, X_2) = Q^*$$

Her problem is to minimize costs, which equal

$$C = X_1P_1 + X_2P_2 \quad (5.22)$$

subject to the constraint that

$$Q^* - f(X_1, X_2) = 0 \quad (5.23)$$

We use Lagrangian multipliers to solve this problem. Again, we first construct the Lagrangian function, which is the right side of equation (5.22) plus λ times the left side of equation (5.23)

$$L_2 = X_1P_1 + X_2P_2 + \lambda[Q^* - f(X_1, X_2)]$$

where λ is the Lagrangian multiplier. Taking the partial derivatives of L_2 with respect to X_1 , X_2 , and λ and setting them all equal to zero, we obtain

$$\frac{\partial L_2}{\partial X_1} = P_1 - \lambda \frac{\partial f(X_1, X_2)}{\partial X_1} = 0 \quad (5.24)$$

$$\frac{\partial L_2}{\partial X_2} = P_2 - \lambda \frac{\partial f(X_1, X_2)}{\partial X_2} = 0 \quad (5.25)$$

$$\frac{\partial L_2}{\partial \lambda} = Q^* - f(X_1, X_2) = 0 \quad (5.26)$$

These are the conditions for cost minimization subject to the output constraint.

Substituting MP_1 for $\partial f(X_1, X_2)/\partial X_1$ and MP_2 for $\partial f(X_1, X_2)/\partial X_2$ in equations (5.24) and (5.25), we get

$$P_1 - \lambda MP_1 = 0$$

$$P_2 - \lambda MP_2 = 0$$

which implies that

$$P_1 = \lambda MP_1 \quad (5.27)$$

$$P_2 = \lambda MP_2 \quad (5.28)$$

Dividing each side of equation (5.27) by the corresponding side of equation (5.28), we find that

$$\frac{P_1}{P_2} = \frac{MP_1}{MP_2}$$

or

$$\frac{MP_1}{P_1} = \frac{MP_2}{P_2}$$

which is our decision rule in equation (5.6).

CHAPTER 6

LEARNING OBJECTIVES

Opportunity Costs
Short-Run Cost Functions
Average and Marginal Costs
Long-Run Cost Functions
Managerial Use of Scale Economies
Managerial Use of Scope Economies
Transactions Costs Can Take Many Forms
Managerial Use of Break-Even Analysis
Profit Contribution Analysis
Summary
Problems
Excel Exercise: Production and Cost
Appendix A: Break-Even Analysis and Operating Leverage
Appendix B: Measurement of Short-Run Cost Functions: The Choice of a Mathematical Form

THE ANALYSIS OF COSTS

Even a manager who fully understands the relationship between inputs and outputs still cannot make optimal (profit-maximizing) decisions without cost information. The key question managers must ponder is this: How are costs related to output? A full understanding of costs is necessary because virtually all business decisions require comparisons of costs and benefits. A manager wants to undertake an action if the additional (marginal) revenue attributable to that action exceeds its additional (marginal) cost. As we will see, to maximize profit, a manager wishes to produce at an output level where the marginal revenue equals the marginal cost. Obviously this calculation is not possible without a knowledge of the cost structure.

Cost (like many four-lettered words) invokes multiple interpretations. Managers find that what seems like a simple concept often provokes controversy over the nature of costs, how they are defined, and their scope and relevance in a decision (hence the basis for cost accounting in virtually every MBA program). A thorough understanding of cost is necessary for a variety of basic managerial decisions: pricing, output, transfer pricing, cost control, and planning for future production.

Managerial consideration of costs must include both short-run and long-run components. A focus on just one of these components, especially the short term, can have catastrophic consequences for an organization. As we detail in later chapters, most managerial decisions require long-term vision.

This chapter explains the basics of cost analysis and describes models to help managers create competitive advantages using cost analysis.

OPPORTUNITY COSTS

Managerial economists define the opportunity cost of producing a particular product as the revenue a manager could have received if she had used her resources to produce the next best alternative product or service. That is, opportunity costs are the revenues forgone if resources (inputs) are not optimally used. They are one reason why managers want to use resources as efficiently as possible; managers need to reduce opportunity costs.

We encounter opportunity costs throughout our lives. Those of you who were accepted at more than one college already have. Those of you who are married or in committed relationships also should understand the concept.

The opportunity cost of General Electric managers' decision to produce large gas turbines is the revenue they could have earned if the labor, equipment, and materials used in the production of turbines were used to produce another GE product—say debt financing. Or GE managers could even have invested outside their firm. Economists believe the true costs of inputs are their values when used in the most productive way. These costs, together with the firm's production costs (the accounting costs of producing a product), determine the economic cost of production. This is called the **opportunity cost doctrine**.

The opportunity cost of an input may not equal its **historical cost**, which is defined as the money managers actually paid for it. For example, if a manager invests \$1 million in equipment that is quickly outmoded and inefficient relative to new equipment, its value is clearly not \$1 million. Although conventional accounting rules place great emphasis on historical costs, managerial economists believe historical costs can be misleading.

Managers must be concerned with two types of costs, both of which are important. The first type is **explicit costs**, which are the ordinary items accountants include as the firm's expenses. These include the firm's payroll, payments for raw materials, and so on. The second type is **implicit costs**, which include the forgone value of resources that managers did not put to their best use (i.e., opportunity costs). Unfortunately accountants and managers, in calculating the costs to a firm, often omit implicit costs.

Think of opportunity costs in the context of MBA students. The total cost (including room and board) of a year's schooling at the Wharton School is roughly \$100,000. This is the cash outlay that most students pay. However, many MBA students held jobs before coming back to school. Assume the compensation of the average MBA student in the previous year was \$70,000. If we asked an accountant the average yearly cost of attending Wharton, he would say about \$100,000. If we pose that same question to economists, most would say \$170,000.

Opportunity cost doctrine The inputs' values (when used in their most productive way) together with production costs (the accounting costs of producing a product) determine the economic cost of production.

Historical cost The money that managers actually paid for an input.

Explicit costs The ordinary items accountants include as the firm's expenses.

Implicit costs The forgone value of resources that managers did not put to their best use.

Consider John Harvey, the proprietor of a firm who invests his own labor and capital in the business. These inputs should be valued at the amount he would have received if he had used them in a different manner. If he could have received a salary of \$65,000 working for someone else and he could have received dividends of \$20,000 by investing his capital elsewhere, he should value his labor and capital at these rates. Excluding these implicit costs can be a serious mistake.

Sunk costs Sunk costs are resources that are spent and cannot be recovered.

Economists also follow the doctrine of **sunk costs**. Sunk costs are resources that are spent and cannot be recovered. For example, if a company builds a plant for \$12 million but then disposes of it for a price of \$4 million, it incurs sunk costs of \$8 million. Sunk costs equal the difference between what a resource costs and what it is sold for in the future.

Ignoring sunk costs is difficult for managers—and in fact perplexes most folks. For example, many people stay in unhealthy relationships because of the time they have invested. You may have given a partner the best five years of your life, but that is no reason to remain in a relationship you don't like. No matter what you do, you cannot recapture the five years, so ignore them in deciding your future.

Rational managers must ignore sunk costs and choose between possible strategies by evaluating only future costs and benefits. For example, if a manager has already spent \$6 million on an advertising campaign, those costs are sunk (they cannot be recovered). So she cannot argue that she has already spent \$6 million and needs only \$1 million more to “turn the corner.” The expected return of that \$1 million in the campaign must be compared to the expected return of \$1 million across alternative investments.

Cost function Function showing various relationships between input costs and output rate.

Short run The time span between one where the quantity of no input is variable and one where the quantities of all inputs are variable.

Fixed inputs When the quantities of plant and equipment cannot be altered.

Scale of plant This scale is determined by fixed inputs.

Variable inputs Inputs that a manager can vary in quantity in the short run.

SHORT-RUN COST FUNCTIONS

Given a firm's cost of producing each level of output, we can define the firm's cost structure. A **cost function** shows various relationships between input costs and output. The firm's production function and the input prices determine the firm's cost structure.

Similar to what we saw with production functions, cost functions are either for the short or long run. The short run is a period so short that a manager cannot alter the quantity of some inputs. As the length of time increases, more inputs become variable. The time span between one where the quantity of no input is variable and one where the quantities of all inputs are variable is called the **short run**. However, a more restrictive definition is generally employed: We say the short run is the time interval so brief that a manager cannot alter the quantities of plant and equipment. These are the firm's **fixed inputs**, and they determine the firm's **scale of plant**. Inputs like labor, which a manager can vary in quantity in the short run, are the firm's **variable inputs**.

TABLE 6.1

Fixed, Variable, and Total Costs: Media Corporation

Units of Output <i>Q</i>	Total Fixed Cost (Dollars per Day) <i>TFC</i>	Total Variable Cost (Dollars per Day) <i>TVC</i>	Total Cost (Dollars per Day) <i>TC</i>
0	100	0	100
1	100	40	140
2	100	64	164
3	100	78	178
4	100	88	188
5	100	100	200
5.5	100	108.625	208.625
6	100	120	220
6.64	100	139.6	239.6
7	100	154	254
8	100	208	308
9	100	288	388
10	100	400	500

We consider three short-run cost concepts: fixed, variable, and total. **Total fixed cost (*TFC*)** is the total cost per period of time incurred for fixed inputs. Because the level of fixed inputs is constant (by definition), the firm's total fixed cost does not vary with output. Examples of fixed costs are depreciation of plant and equipment and property taxes. Table 6.1 shows that the fixed cost of the Media Corporation, a producer of sofas, is \$100 per day. This is visually shown in Figure 6.1. If *Q* is equal to total output, the values in Table 6.1 (and in Table 6.2) come from the total cost relationship

$$TC = 100 + 50Q - 11Q^2 + Q^3$$

Total variable cost (*TVC*) is the total cost incurred by managers for variable inputs. These costs increase as output rises because greater output requires more inputs and higher variable costs. For example, the greater the output of a woolen mill, the larger the quantity of wool used and the higher the total cost of the wool. The Media Corporation's total variable cost schedule is shown in Table 6.1. Figure 6.1 shows the corresponding total variable cost function. Up to a particular output rate (four units of output), total variable costs rise at a decreasing rate; beyond that output level, they increase at an increasing rate. This characteristic of the total variable cost function follows from the law of diminishing marginal

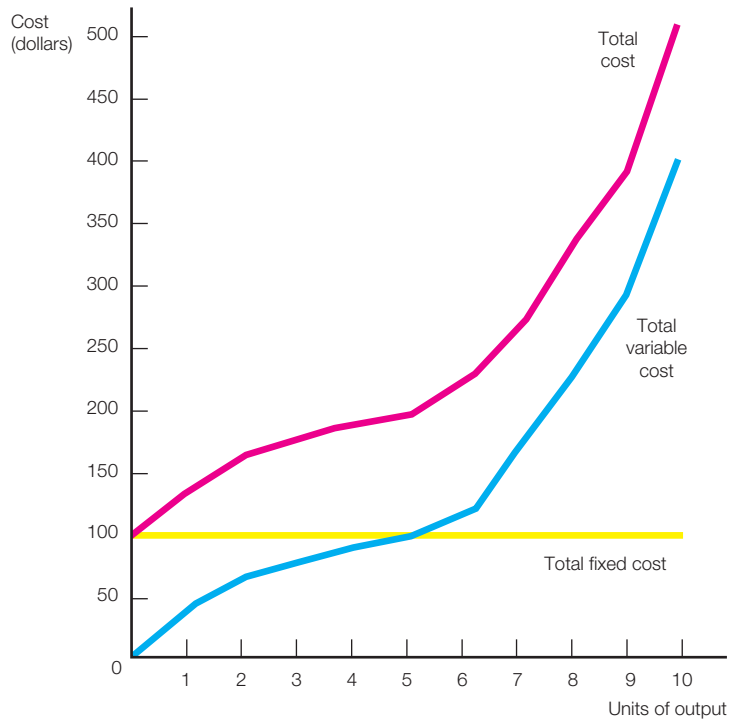
Total fixed cost (*TFC*) The total cost per period of time incurred for fixed inputs.

Total variable cost (*TVC*) The total cost incurred by managers for variable inputs.

FIGURE 6.1

Fixed, Variable, and Total Costs: Media Corporation

Fixed costs do not vary with output, so the fixed cost curve is a horizontal line. Variable costs at first increase with output at a decreasing rate and then increase with output at an increasing rate. The total cost curve is the vertical summation of the fixed cost curve and the average variable cost curve. The total cost function and the total variable cost function have the same shape because they differ by only a constant amount, which is total fixed cost.



returns. At low levels of output, increasing the variable inputs may increase productivity, with the result that total variable costs rise with output but at a decreasing rate. (More will be said about this later.)

Total cost (TC) The sum of total fixed and total variable costs.

Finally, **total cost (TC)** is the sum of total fixed and total variable costs. To derive the total cost column in Table 6.1, add the total fixed cost and total variable cost at each output level. The total cost function for the Media Corporation is also shown in Figure 6.1. The total cost function and the total variable cost function have the same shape because they differ by only a constant amount, which is total fixed cost.

Managers want to allocate resources efficiently (remember, they are scarce). They want to choose the input bundle that produces a given output at the lowest

possible cost. Costs are a function of the production process and input prices, so managers need to understand the behavior of costs as output changes. One indicator of this behavior is found in average and marginal costs.

AVERAGE AND MARGINAL COSTS

Although the total cost of producing a product is important, from an operational viewpoint, knowledge of the average and marginal cost functions is key. These functions predict the behavior of costs as output changes. There are three average cost functions, corresponding to the three total cost functions. **Average fixed cost (AFC)** is total fixed cost divided by output. *AFC* necessarily declines with increases in output; mathematically, the function is a rectangular hyperbola. Table 6.2 and Figure 6.2 show the *AFC* function for the Media Corporation.

Average variable cost (AVC) is total variable cost divided by output. This indicator tells managers the variable cost, on average, of each unit of output. For the Media Corporation, the *AVC* function is shown in Table 6.2 and Figure 6.2. Initially,

Average fixed cost (AFC) The total fixed cost divided by output.

Average variable cost (AVC)
The total variable cost divided by output.

TABLE 6.2

Average and Marginal Costs: Media Corporation

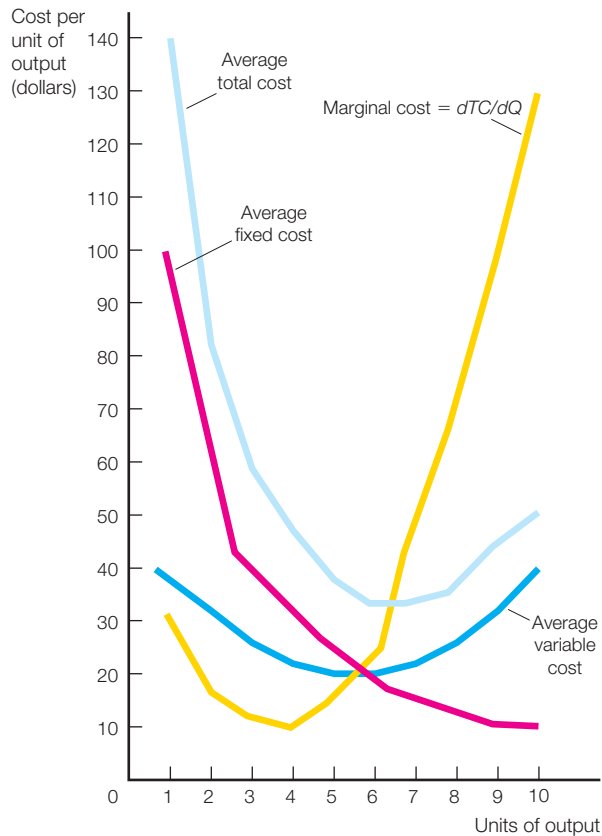
Units of Output <i>Q</i>	Average Fixed Cost (Dollars) <i>TFC/Q</i>	Average Variable Cost (Dollars) <i>TVC/Q</i>	Average Total Cost (Dollars) <i>TC/Q</i>	Marginal Cost (Dollars) $\Delta TC/\Delta Q^a$	Marginal Cost (Dollars) dTC/dQ^a
0	—	—	—	—	—
1	100	40	140	40	31
2	50	32	82	24	18
3	33.33	26	59.33	14	11
4	25	22	47	10	10
5	20	20	40	12	15
5.5	18.18	<u>19.75</u>	37.93		<u>19.75</u>
6	16.67	20	36.67	20	26
6.64	15.06	21.04	<u>36.11</u>		<u>36.11</u>
7	14.29	22	36.29	34	43
8	12.5	26	38.5	54	66
9	11.11	32	43.11	80	95
10	10	40	50	112	130

^a The figures in the $\Delta TC/\Delta Q$ column pertain to the interval between the indicated amount of quantity and one unit less than the indicated amount of quantity. The figures in the dTC/dQ column are the continuous marginal cost—that is $dTC/dQ = MC = 50 - 22Q + 3Q^2$.

FIGURE 6.2

Average and Marginal Cost Curves: Media Corporation

Average fixed cost continually decreases as output increases. Average variable cost and average total cost at first decrease, reach a minimum, then increase as output increases. The minimum of the average total cost occurs at a higher output than the minimum of the average variable cost. The average total cost curve is the vertical summation of the average fixed cost and the average variable cost curves. Marginal cost passes through the minimum of both average cost curves, and when marginal cost is below the average cost, average cost falls and vice versa. Average total cost achieves its minimum at a higher output rate (6.64) than average variable cost (5.5) because the increases in average variable cost are, up to a point, more than offset by decreases in average fixed cost.



increasing output results in decreases in *AVC*. However, as output increases, at some point of increased production, *AVC* rises, thus increasing the average variable cost per unit.

Last chapter we said the behavior of cost is largely determined by the production function. Now we can show why. *AVC* is simply total variable cost divided by

the number of units produced (TVC/Q). Let U be the number of input units used and W the cost per unit of input. It is then true that

$$AVC = \frac{TVC}{Q} = W \frac{U}{Q}$$

We defined average product in Chapter 5 as

$$AP = \frac{Q}{U}$$

Hence AP is the inverse of U/Q . Now we can express AVC as

$$AVC = W \frac{1}{AP} \quad (6.1)$$

So one way for managers to think of AVC is as the inverse of average product times the cost per unit of input. Recall the behavior of the average product. It initially increases with output, reaches a maximum, and then begins to decrease. Because of their inverse relationship, AVC mirrors the behavior of AP . When AP increases, AVC decreases; when AP decreases, AVC increases. So we expect AVC to initially decrease, hit a minimum, and then increase.

Average total cost (ATC) is total cost divided by output. For the Media Corporation, the ATC function is shown in Table 6.2 and Figure 6.2. ATC is simply the sum of $AFC + AVC$. Its shape is similar to AVC but higher at all output levels due to the effect of fixed costs. At output levels where both AFC and AVC decrease, ATC must decrease too. However, ATC reaches its minimum at output levels relatively higher than AVC because increases in average variable cost are for a time more than offset by decreases in average fixed cost (which must decrease as output increases).

Average total cost (ATC) The total cost divided by output.

Marginal cost (MC) is the incremental cost of producing an additional unit of output. If $C(Q)$ is the total cost of producing Q units of output, the marginal cost between Q and $(Q - 1)$ units of output is $C(Q) - C(Q - 1)$. For the Media Corporation, the marginal cost function is shown in Table 6.2 and Figure 6.2. At low output levels, MC may decrease (as it does in Figure 6.2) with increases in output; but after reaching a minimum, it increases (like AVC) with additional output. We saw why this is true last chapter when we discussed diminishing marginal returns. If ΔTVC is the change in total variable costs resulting from a change in output of ΔQ and if ΔTFC is the change in total fixed costs resulting from a change in output of ΔQ then

Marginal cost (MC) The incremental cost of producing an additional unit of output.

$$MC = \frac{\Delta TVC + \Delta TFC}{\Delta Q}$$

But ΔTFC is zero because fixed costs can't vary; therefore

$$MC = \frac{\Delta TVC}{\Delta Q}$$

Hence the cost of the input is given by $\Delta TVC = W(\Delta U)$, where W is the cost per unit of input and ΔU is the change in the units of input needed to produce the increase of ΔQ in output. Consequently,

$$MC = W \frac{\Delta U}{\Delta Q}$$

Last chapter we defined MP as

$$MP = \frac{\Delta Q}{\Delta L}$$

Hence we can define MC as

$$MC = W \frac{1}{MP}$$

So like the inverse nature of AP and AVC , the behavior of MP is inverse to that of MC . Marginal cost is simply the cost per unit of input times the inverse of its marginal product. Let's think about why this is true. Say a unit of labor costs \$10. If the MP of a unit of labor is 10, the MC of producing the last unit of output is \$1 (\$10/10). But if the MP of that unit of labor is 1, the MC is \$10. As MP increases, MC decreases; and when MP decreases, MC increases. We saw that the behavior of marginal product is to increase, attain a maximum, and then decline with increases in output; marginal cost normally decreases, attains a minimum, and then increases.

QUANT OPTION

If the total cost function is continuous, marginal cost is defined as dTC/dQ , where TC is total cost. Suppose, for example, using the Media Corporation's total cost function,

$$TC = 100 + 50Q - 11Q^2 + Q^3$$

where TC is expressed in thousands of dollars and Q is expressed in units of output. This firm's marginal cost function is

$$MC = \frac{dTC}{dQ} = 50 - 22Q + 3Q^2$$

TABLE 6.3

Relationship of Average Product and Marginal Product to Average Variable Cost and Marginal Cost: Thomas Machine Company

L	Q	AP_L	$MP_L = dQ/dL$	W	$AVC = W/AP_L$	$MC = W/MP_L$
0	0	—	—	390	—	—
1	49	49	67	390	7.96	5.82
2	132	66	98	390	5.91	3.98
3	243	81	123	390	4.81	3.17
4	376	94	142	390	4.15	2.75
5	525	105	155	390	3.71	2.52
6	684	114	162	390	3.42	2.41
6.67	792.6	118.9	163.33	390	3.28	2.388 $\leftarrow MP_L$ max
7	847	121	163	390	3.22	2.393 so MC min
8	1008	126	158	390	3.10	2.47
9	1161	129	147	390	3.02	2.65
10	1300	130	130	390	3.00 $\leftarrow AP_L$ max	3.00
11	1419	129	107	390	3.02 so AVC min	3.64
12	1512	126	78	390	3.10	5.00
13	1573	121	43	390	3.22	9.07
14	1596	114	2	390	3.42	195.00
15	1575	105	-45	390	3.71	—

The relationship between production and cost is shown in Table 6.3. Consider the previous definitions of the production function of the Thomas Machine Company ($Q = 30L + 20L^2 - L^3$) from the last chapter. If the wage rate is 390, Table 6.3 shows the relationship between average product and average variable cost and between marginal product and marginal cost. As you can see, when AP is maximized, AVC is minimized. Likewise, we see that MC equals AVC when average variable cost is minimized. This is to be expected because $MP = AP$ when AP is maximized. Average product is maximized when 10 units of labor are employed (at 130); average variable cost is minimized when 10 units of labor are employed (at 3); and marginal cost is also equal to 3 when 10 units of labor are employed.

Marginal cost always equals average variable cost when the latter is at a minimum (because $MP = AP$, when AP is maximized). If the cost function of Thomas is $TC = 100 + 50Q - 11Q^2 + Q^3$, then the firm's AVC is

$$AVC = \frac{TVC}{Q} = 50 - 11Q + Q^2$$

If we take the ΔAVC with respect to ΔQ and set it equal to zero, we find the value of Q where AVC is at a minimum

$$\frac{\Delta AVC}{\Delta Q} = -11 + 2Q = 0$$

$$Q = 5.5$$

When Q equals 5.5, both marginal cost and average variable cost equal \$19.75. (Substitute 5.5 for Q in the preceding equations for MC and AVC and see for yourself that this is true.) Therefore, as pointed out, $MC = AVC$ when AVC is at a minimum. Note also that marginal cost equals average total cost when the latter is at a minimum. The firm's average total cost is

$$ATC = (100/Q) + 50 - 11Q + Q^2$$

If we take ΔATC with respect to ΔQ and set it equal to zero, we find the value of Q where ATC is at a minimum

$$\frac{\Delta ATC}{\Delta Q} = \left(\frac{-100}{Q^2} \right) - 11 + 2Q = 0$$

Or

$$2Q^3 - 11Q^2 - 100 = 0$$

This is solved for $Q = 6.64$. Substituting 6.64 into the ATC and MC equations yields $MC = ATC = 36.11$.

QUANT OPTION

Here's some fun!

$$\begin{aligned} \frac{dAVC}{dQ} &= \frac{d\left(\frac{VC}{Q}\right)}{dQ} \\ &= \frac{\left[Q\left(\frac{dVC}{dQ}\right) - VC\left(\frac{dQ}{dQ}\right)\right]}{Q^2} \\ &= \frac{\left[\left(\frac{dVC}{dQ}\right) - \left(\frac{VC}{Q}\right)\right]}{Q} = 0 \end{aligned}$$

which implies

$$\left(\frac{dVC}{dQ}\right) - \left(\frac{VC}{Q}\right) = MC - AVC = 0$$

Or

$$MC = AVC$$

PROBLEM SOLVED: The Effects of Output on the Cost of Producing Aircraft

The National Research Council conducted a study of the U.S. aircraft industry, which stresses the importance to airplane manufacturers of serving the entire world market. As evidence, the council presents the following graph.

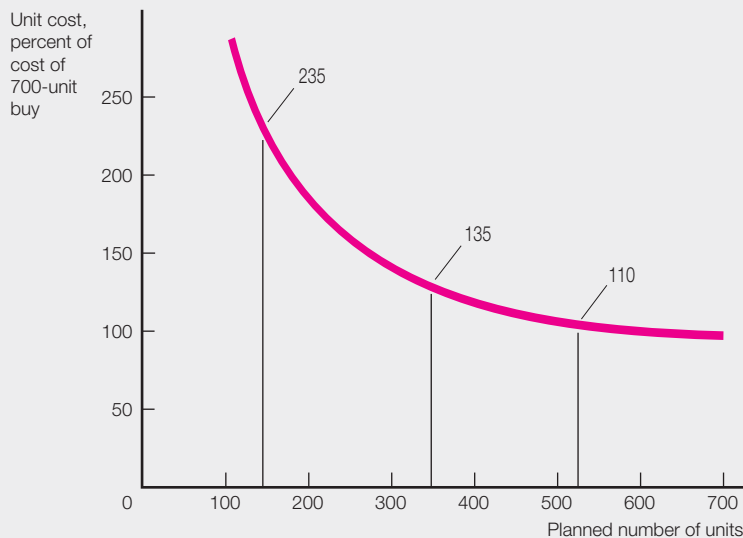
Problems

1. As indicated in this graph, the cost per airplane of producing 525 aircraft of a particular type is about 10% higher than the cost per airplane of producing 700 aircraft of this type. Assuming this graph pertains to the short run, by what percentage does average fixed cost increase if 525 rather than 700 aircraft are produced?
2. If average fixed cost is 30% of average total cost if 700 aircraft are produced and 36% of average total cost if 525 aircraft are produced, is it true that aver-

age total cost is about 10% higher if 525 rather than 700 aircraft are produced?

Solutions

1. If the number of aircraft produced is 525 rather than 700, average fixed cost is $TFC/525$ rather than $TFC/700$, where TFC equals total fixed cost. Therefore, average fixed cost increases by 33%.
2. For 700 aircraft, average total cost equals $X/0.30 = 3.33X$, where X is average fixed cost when 700 aircraft are produced. For 525 aircraft, average total cost equals $1.33X/0.36 = 3.69X$ because average fixed cost equals $1.33X$ when 525 aircraft are produced. Therefore, average total cost increases by about 11% (from $3.33X$ to $3.69X$) if 525 rather than 700 aircraft are produced.



Note that we define marginal cost in two ways. The first ($\Delta TC/\Delta Q$) assumes we can produce only in discrete units, like a car or a cake. The second (dTC/dQ) assumes we can produce on a continuous basis, as in 3.14 tons of grain or 10.33 gallons of gasoline. Which definition managers use depends on the situation: Is the manager constrained to an integer output level or can she produce fractions of the product?

LONG-RUN COST FUNCTIONS

In the long run all inputs are variable, and managers can build any scale or type of plant. There are no fixed costs in the long run because no inputs are fixed; so there is nothing to stop a manager from being as efficient as possible. Thinking about the long run requires managers to focus more on the destination rather than the route. As we show in Chapters 11 and 12, managers need to actively anticipate the future and think toward it.

When Toyota managers consider building a new U.S. plant, they can build one of any size at any location; there are boundless choices. But once the investment is made, type and size of plant and equipment are to a considerable extent frozen.

Assume managers are looking to construct one of three alternative scales of plant; the short-run average cost functions for these scales of plant are represented by G_1G_1 , G_2G_2 , and G_3G_3 , in Figure 6.3. In the long run managers can choose any of the three plants. Which scale is most profitable? Obviously the answer depends on the manager's beliefs about long-run product demand because she needs to produce at the minimum average cost.

We can see this in Figure 6.3. If the manager anticipates product demand of around Q , she should build the smallest plant. At these sales, the average cost per unit sold is equal to C . If the manager builds the medium-sized or largest plant, the average cost per unit is higher at Q . If the manager believes that demand will be equal to S , she should build the largest plant (G_3).

The **long-run average cost function** (*LAC*) shows the minimum cost per unit of all output levels when any desired size plant is built. In Figure 6.3 the long-run average cost function is the solid portion of the short-run average cost functions, $G_1DEG'_3$. Any point on the long-run average cost function is also a point on a short-run average cost function. In fact, it is a point on the lowest-cost short-run cost function for the given output level. So when given the freedom (that is, in the long run), managers want to choose the plant scale that minimizes average cost. The broken-line segments of the short-run functions are not included because they are not the lowest average costs, as is obvious from the figure.

Toyota managers can choose from more than three plant scales: They can choose from an infinite number of possibilities. However, managers must understand that once they commit funds to building a plant, they immediately shift to a short-run cost function. Figure 6.4 depicts this decision. The minimum average

Long-run average cost function (LAC) Function showing the minimum cost per unit of all output levels when any desired size plant is built.

FIGURE 6.3

Short-Run Average Cost Functions for Various Scales of Plant

The long-run average cost function is the solid portion of the short-run average cost functions, $G_1 DEG'_3$.

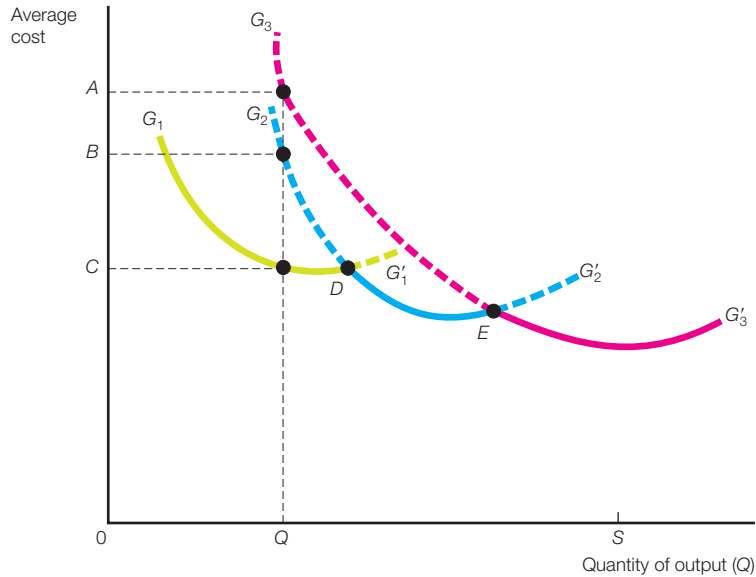
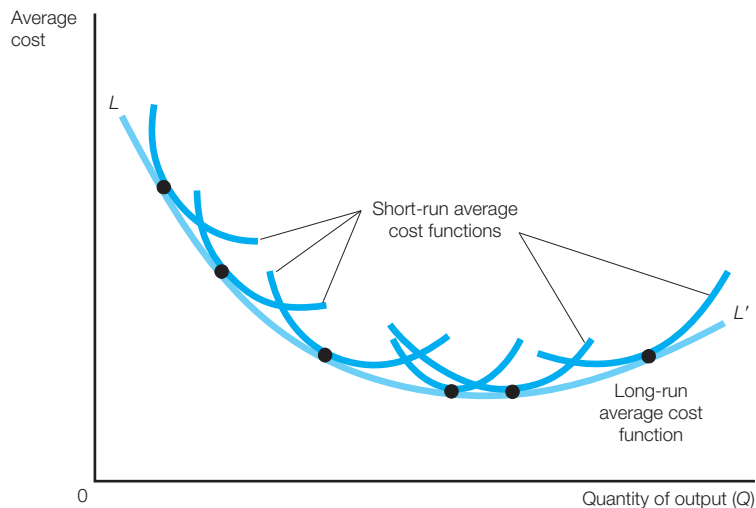


FIGURE 6.4

Long-Run Average Cost Function

The long-run average cost function, which shows the minimum long-run cost per unit of producing each output level, is the envelope of the short-run functions.



PROBLEM SOLVED: Managerial Use of Cost Functions

To illustrate the relationship between a firm's long-run and short-run cost functions, consider a Martin division that produces MP3 players. Managers have been told to cut costs, so they need to create an efficient input bundle. Engineers estimate the division's production function as

$$Q = 4(KL)^{0.5} \quad (6.2)$$

where Q is output (in thousands of MP3 players per month), K is the capital used per month (in thousands of units), and L is the number of hours of labor employed per month (in thousands). Because laborers are paid \$8 per hour and capital costs equal \$2 per unit, Martin's total cost (in thousands of dollars per month) equals

$$TC = 8L + 2K = \frac{Q^2}{2K} + 2K \quad (6.3)$$

Equation (6.2) implies that

$$L = \frac{Q^2}{16K}$$

In the short run, managers cannot vary their level of capital, so K is fixed. The division presently uses 10,000 units of capital ($K = 10$). Substituting 10 for K in equation (6.3), we find that the short-run cost function is

$$TC_S = \frac{Q^2}{20} + 20 \quad (6.4)$$

where TC_S is short-run total cost. Therefore, the short-run average total cost function is

$$AC_S = \frac{TC_S}{Q} = \frac{Q}{20} + \frac{20}{Q}$$

and the short-run marginal cost function is

$$MC_S = \frac{\Delta TC_S}{\Delta Q} = \frac{Q}{10}$$

In the long run, managers can buy new machines or sell current ones, so no input is fixed. If Martin managers want to minimize total costs, they need to choose the optimal level of capital needed to produce Q MP3 players. Basing their estimates on equation (6.3), they find that

$$\frac{\Delta TC}{\Delta K} = -\frac{Q^2}{2K^2} + 2$$

Setting this equal to zero, we find that the cost-minimizing value of K is

$$K = \frac{Q}{2}$$

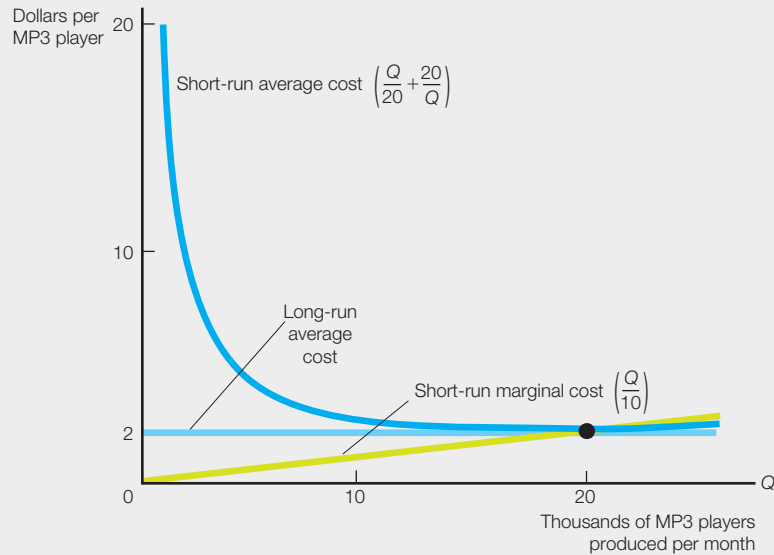
This decision rule tells the managers to estimate expected demand and then purchase capital equal to one-half of demand. Substituting $Q/2$ for K in equation (6.3), we see that the long-run cost function is

$$TC_L = 2Q \quad (6.5)$$

where TC_L is long-run total cost. Because $TC_L/Q = 2$, the long-run average cost equals \$2 per MP3 player. The long-run marginal cost is also \$2 for MP3 players because $\Delta TC_L/\Delta Q = 2$.

Short-Run Average and Marginal Costs and Long-Run Average Cost, Martin Division

Because the long-run average cost function is horizontal, it is tangent to the short-run average cost function at the latter's minimum point.



The figure shows the relationship between Martin's short-run average and marginal costs and its long-run average costs. As managers expect, the short-run marginal cost function intersects the short-run average cost function at its minimum point, where $Q = 20$ and $AC_S = 2$, in this case. Because it is horizontal (owing to constant returns to scale), the long-run average cost function is tangent to the short-run average

cost function at the latter's minimum point. Many long-run average cost functions are not horizontal. Instead they have what we call *scale economies*: As the plant size varies, so does the average unit cost. As discussed in the following section, there are economies of scale (over at least some range of output) in a wide variety of markets and processes. Smart managers use these scale economies to create competitive advantage.

cost of producing all outputs is given by the long-run AC function, LL' . Each point on LL' is also a point on a short-run AC function. At that output level, it is a point on the lowest-cost short-run AC function; it is the best an efficient manager can do. The two functions are tangent at that point. (Mathematically, the long-run average cost function is the envelope of the short-run functions.)

Managers who estimate the long-run average cost of producing any given output can readily derive the long-run total cost of production: It is simply the product of long-run average cost and output. Figure 6.5 shows the relationship between long-run total cost and output; this relationship is called the **long-run total cost function**.

Managers can readily derive the **long-run marginal cost function**. This function represents how varying output affects the cost of producing the last unit if the manager has chosen the most efficient input bundle. This marginal cost function shows behavior similar to average costs. Long-run marginal cost is less than LAC when LAC is decreasing; it is equal to long-run LAC when LAC is at a minimum; and it is greater than LAC when LAC is increasing. When managers build the optimal scale of plant for producing a given level of output, long-run marginal cost is equal to short-run marginal cost at that output.

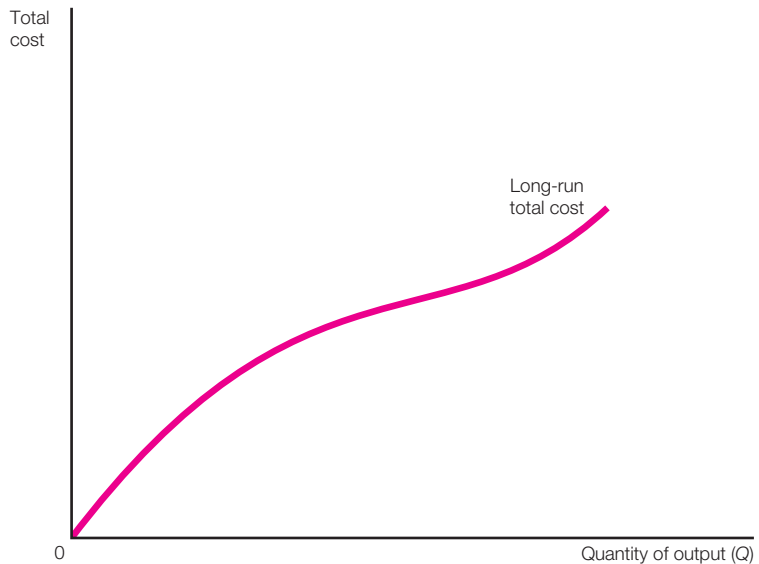
Long-run total cost function The relationship between long-run total cost and output.

Long-run marginal cost function Function representing how varying output affects the cost of producing the last unit if the manager has chosen the most efficient input bundle.

FIGURE 6.5

Long-Run Total Cost Function

The long-run total cost of a given output level equals the long-run average cost (given in Figure 6.4) times output.



QUANT OPTION

Never believe everything you read; always ask for the proof.

Suppose the long-run average cost of producing an output rate of Q is $L(Q)$, and the short-run AC of producing it with the i th scale of plant is $A_i(Q)$. Let $M(Q)$ be the long-run MC and $R_i(Q)$ be the short-run MC with the i th scale of plant. If the manager maximizes profit, she is operating where short-run and long-run average costs are equal; in other words, $L(Q) = A_i(Q)$. This means that

$$\frac{dL(Q)}{dQ} = \frac{dA_i(Q)}{dQ} \text{ and } Q \frac{dL(Q)}{dQ} = Q \frac{dA_i(Q)}{dQ}$$

From these conditions, it is easy to prove that the long-run marginal cost, $M(Q)$, equals the short-run marginal cost, $R_i(Q)$

$$M(Q) = \frac{d[QL(Q)]}{dQ} = L(Q) + \frac{QdL(Q)}{dQ}$$

$$R_i(Q) = \frac{d[QA_i(Q)]}{dQ} = A_i(Q) + \frac{QdA_i(Q)}{dQ}$$

We know from the previous paragraph that $L(Q) = A_i(Q)$ and $QdL(Q)/dQ = QdA_i(Q)/dQ$; so it follows that $R_i(Q)$ must equal $M(Q)$.

QUANT OPTION

$$dTC_S/dQ = Q/10$$

$$dTC/dK = -(Q^2/2K^2) + 2 = 0$$

$$dTC_L/dQ = 2$$

MANAGERIAL USE OF SCALE ECONOMIES

Long-run average cost curves tell managers whether bigger is better: They show whether, and to what extent, larger plants have cost advantages over smaller ones.

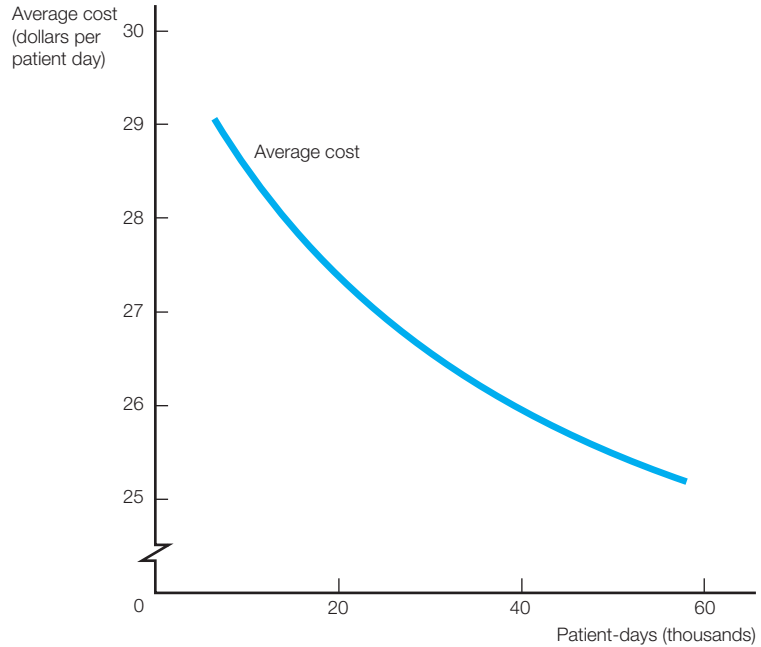
Economies of scale occur when the firm's average unit cost decreases as output increases. To illustrate, consider nursing homes, which make up an industry with

Economies of scale When the firm's average unit cost decreases as output increases.

FIGURE 6.6

Long-Run Average Cost Curve for Texas Nursing Homes

For nursing homes with fewer than 60,000 patient-days, there seem to be substantial economies of scale.



annual sales of over \$190 billion in 2011. Figure 6.6 shows the long-run average cost curve for a nursing home.

As you can see, there are substantial economies of scale. If a nursing home provides 10,000 patient-days of service per year, the cost per patient-day is almost \$29; if it provides about 50,000 patient-days of service per year, the cost per patient-day is under \$26. Curves like Figure 6.6 are estimated by engineers and economists for a wide variety of plants and processes; they help managers choose optimal input bundles.

Scale economies are not confined to plants. In our global economy, managers often use more than one plant to produce their product. Nike is a good example. Managers at such companies have the opportunity to exploit possible scale economies at the firm level. For example, holding the size of each plant constant, average cost may decrease in response to increases in the number of plants operated by Nike. Managers could create cost savings for several reasons, including centralized purchasing, better management techniques, or an improved ability to build or lease plants.

Managers use many sources of scale economies to create competitive advantages. Managers at UPS use them in their distribution network to decrease costs. Exxon Mobil managers use them to decrease costs in their refining and chemical process. Managers at BskyB, the British pay television server, can pay more for some content because of scale economies in its large network. Finally, the size of cruise ships keeps growing because larger ships have a lower cost per passenger thanks to scale economies. Royal Caribbean continues this trend with the launching of the *Oasis of the Seas* ship in 2010; this 225,000-ton ship can carry 6,296 passengers. When asked why bigger is better, Harri Kulovaara of Royal Caribbean noted, “Having more real estate, we can provide more deck area. That means more entertainment options and better amenities.”¹

Managers must understand their cost relationships to recognize where to best exploit scale economies. As we have seen, scale economies are not confined to production; they are found in distribution, raising capital, advertising, and most business processes. All managers have the opportunity to exploit scale economies in some form, though some fail to recognize their opportunities.

However, bigger is not always better. As plants, distribution networks, or cruise ships get bigger, at some point managing them gets harder. Increasing size eventually causes **diseconomies of scale**: Average costs per unit of output increase, usually because of the complexity of managing and coordinating all the necessary activities.

Diseconomies of scale When the average costs per unit of output increase.

MANAGERIAL USE OF SCOPE ECONOMIES

Scale economies are not the only cost economies managers can exploit. A cost efficiency strategy available to managers of multiproduct firms is called **economies of scope**. These economies exist when the cost of producing two (or more) products jointly is less than the cost of producing each one alone. For example, suppose managers of the Martin Company can produce 1,000 milling machines and 500 lathes per year at a cost of \$15 million, whereas if the firm produced only 1,000 milling machines, the cost would be \$12 million, and if it produced only 500 lathes, the cost would be \$6 million. In this case the cost of producing both the milling machines and the lathes is less than the total cost of producing each separately. Hence there are scope economies.

A simple way for managers to estimate the extent of their scope economies is to use the following measure

$$S = \frac{C(Q_1) + C(Q_2) - C(Q_1 + Q_2)}{C(Q_1 + Q_2)} \quad (6.6)$$

Here S is the degree of economies of scope, $C(Q_1)$ is the cost of producing Q_1 units of the first product alone, $C(Q_2)$ is the cost of producing Q_2 units of the second product alone, and $C(Q_1 + Q_2)$ is the cost of producing Q_1 units of the

Economies of scope Exist when the cost of jointly producing two (or more) products is less than the cost of producing each one alone.

1. J. Wise, “World’s Largest Cruise Ship Pulls 360s with Joystick,” *Popular Mechanics*, June 2007.

first product in combination with Q_2 units of the second product. If there are economies of scope, S is greater than zero because the cost of producing both products together— $C(Q_1 + Q_2)$ —is less than the cost of producing each alone— $C(Q_1) + C(Q_2)$. Clearly S measures the percentage of saving that results from producing them jointly rather than individually. Managers of the Martin Company calculate the following

$$S = \frac{\$12 \text{ million} + \$6 \text{ million} - \$15 \text{ million}}{\$15 \text{ million}} = 0.20$$

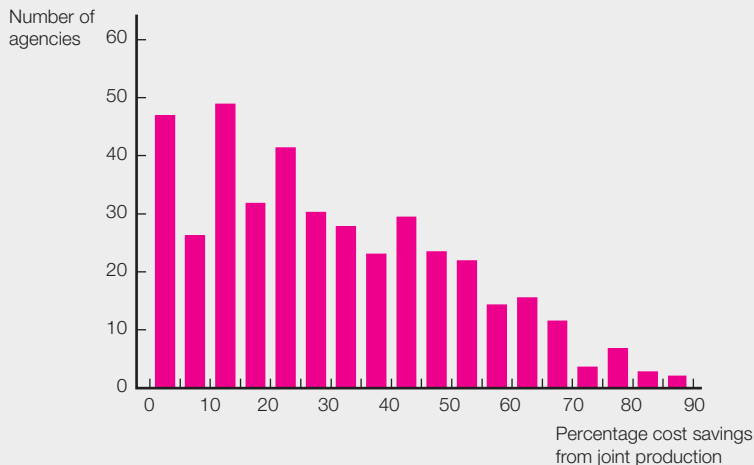
which means scope economies have lowered their costs by an estimated 20%. Strategically, managers of Martin have created a 20% cost advantage relative to managers of single-product firms. Obviously, the larger is the value of S , the greater are the scope economies.

STRATEGY SESSION: Economies of Scope in Advertising Agencies

In recent years there has been considerable controversy over the extent to which there are economies of scope in the advertising industry. An advertising agency can use many media, including network television, general magazines, newspapers, radio, outdoor ads, the Internet, and cell phones.

Researchers have looked at the percentage of cost reduction from joint production at several hundred advertising agencies and found the following cost savings.

As you can see, the cost savings from joint production of these products range from essentially zero to about 86%, depending on which advertising agency is considered. On the average, the cost saving is about 26%. Clearly advertising has very substantial economies of scope; and smaller firms seem to enjoy greater scope economies than larger firms.



STRATEGY SESSION: Economies of Scope

First we had individual firms, then conglomerates, and now perhaps a trend of returning to individual firms focusing on core competencies.

The argument for conglomerates were those of economies of scope—that is, one could produce X and Y cheaper together, $C(X, Y)$, than producing X and Y separately, $C(X) + C(Y)$, on the supply side. Conglomerates also serve as a portfolio diversified “mutual fund” for investors on the demand side. The leader of conglomerates is ITT, which had over 300 separate businesses in the 1960s.

But some have argued that the emperor had no clothes, that conglomerates had overpaid (acquisitions are usually at a premium), and that managers often knew little of the intricacies of the businesses they had acquired. Economies of scope didn’t materialize in many conglomerates and, in fact, evidence of diseconomies existed—does Daimler-Chrysler ring a bell? If investors wanted to diversify their portfolios, they could make their own portfolio by purchasing the stocks of separate companies or buying existing mutual funds.

Recently ITT, Fortune Brands, Marathon Oil, and Cargill, among others, have announced breakups of their companies. These would be what are called *spin-*

offs. In a spinoff, the base company (say ITT) issues shares specifically for the division they wish to spinoff and distributes them on a pro rata basis to their base shareholders. These shareholders can hold the shares or sell them. Spinoffs are used, for example, when the base company can’t sell the subsidiary. Sometimes the base company loads the spinoff with weak portions of the base company’s management and debt. Not a recipe for success of the spinoff.

Some conglomerates, such as General Electric, Berkshire Hathaway, and United Technologies, have been historically successful, constantly reinventing themselves by adding new subsidiaries and shedding old ones, but always with the goal of economies of scope and risk diversifying and not growth for growth’s sake. Some recent acquisitions, such as Disney’s acquisition of Marvel and eBay’s acquisition of PayPal, would seem to make sense from a scope perspective and because of their demand complementarities.

Source: Steven M. Davidoff, “A Test to See If the Parts Are Worth More Than the Whole,” *New York Times*, January 18, 2011, at <http://dealbook.nytimes.com/2011/1/18>.

Like scale economies, managers can exploit scope economies across a range of markets, processes, and behavior. Oil firms like Exxon Mobil and British Petroleum produce both petroleum and chemical products; drug firms like Merck and GlaxoSmithKline produce both vaccines and tranquilizers; and publisher W. W. Norton publishes both economics and literature textbooks. Both Coca-Cola and Pepsi use single trucks to deliver an assortment of flavors and sizes of drinks. Most airlines fly both passengers and cargo. Managers use scope economies to create cost advantages by producing multiple products rather than just one. Often these cost savings arise because the products share either processes (like distribution) or resources (such as components). However, just as with diseconomies of scale, there can be diseconomies of scope. Many conglomerates, such as General

STRATEGY SESSION: Corporate Synergy in Conglomerates

Why do some firms organize as conglomerates whereas others produce in just one line of business? The latter argue for “core competency”—doing what they do best, whereas the former argue for economies of scope and diversification of risk. Those who have disassembled conglomerates argue that the whole is worth less than the sum of its parts (for example, Centent, which was built and then dismantled by Henry Silverman). Those who favor con-

glomerates, such as General Electric and Siemens, argue the reverse and point to synergies of idea generation and the ability to use personnel across businesses.

On the risk-diversification side, critics argue that as capital markets have become more sophisticated, global, and liquid, fund managers maintain that they can diversify risk and increase returns by purchasing securities across multiple sectors.

Electric, continue to subtract (and add) from their lines of business when it is no longer profitable for them to produce a particular product.

TRANSACTIONS COSTS CAN TAKE MANY FORMS

Consider our consumer, Ms. Popovich. She may not know the price of the unit of food in Chapter 3 precisely. But she may know that it typically ranges from P_{Fmin} to P_{Fmax} . Suppose that she currently finds the item at P_{F2} , near the high of P_{Fmax} . She may wish to continue to search for the item knowing that it's likely that she will find it for a lower price. However, searching takes time (even online) and may involve transportation costs (traveling among potential suppliers). What should Ms. Popovich do? She should consider the expected benefits of more search (a lower price) versus the costs of the search (time and money).

The same analysis would apply for a manager faced with input costs that vary. Consider a corporate recruiter. Should she interview at schools X, Y, and B, or just at X? If her experience at school X has been that she most often finds the right person for the job there, she may not visit the other schools. But if her experience is that sometimes X yields the best person and sometimes Y does, she may want to schedule both schools. And what about up and coming school B? Of course, the more schools she visits, the larger the costs, such as the opportunity costs of work she would be doing for her firm if not recruiting—travel, and lodging costs, and so on. The authors are always amazed at the number of management personnel from investment banks and consulting firms that come to Philadelphia for a corporate show and tell. As with Ms. Popovich, the recruiter must consider the expected benefits of more search with the costs of the search.

Another transaction cost is negotiation. Consider our case of the Haddonfield Brewery and the Cherry Hill Chemical Company in Chapter 17. In this example, *in the absence of negotiation costs*, the cheapest (and societally optimal) way to solve the problem was for Haddonfield to pay Cherry Hill to filter its own effluent if Cherry Hill is not liable for the damages its effluent causes or for Cherry Hill to filter its own effluent if Cherry Hill is liable for its damages.

But suppose that Cherry Hill is not liable, and negotiation costs are \$10,001, with Cherry Hill's cost N_C and Haddonfield's costs N_H , where $N_C + N_H = 10,001$ or $N_H = 10,001 - N_C$. Recall that Cherry Hill can filter its outgoing water for a cost of \$40,000, whereas Haddonfield can filter its incoming water for a cost of \$50,000.

Since Cherry Hill is not liable, it is entitled to make its full profit of \$500,000. In order for Cherry Hill to filter its outgoing water, Haddonfield must pay Cherry Hill at least the costs (\$40,000) of Cherry Hill's filtering (since Cherry Hill is not obligated to do so). Call this payment P . Clearly, Haddonfield won't pay Cherry Hill more than \$50,000 (since Haddonfield could filter the incoming water themselves for \$50,000). With Haddonfield currently filtering its incoming water for \$50,000, Haddonfield's profits are \$200,000. For Haddonfield to be willing to negotiate, they must make more than their current profit of \$200,000.

Thus for Cherry Hill to be willing to negotiate with Haddonfield, the following condition must hold

$$\$500,000 - \$40,000 + P - N_C \geq \$500,000$$

Or
$$P - N_C \geq \$40,000$$

For Haddonfield to be willing to negotiate with Cherry Hill, the following condition must hold

$$\$250,000 - P - N_H \geq \$200,000$$

Or
$$\$250,000 - P - (\$10,000 - N_C) \geq \$200,000$$

Or
$$\$240,000 - P + N_C \geq \$200,000$$

Or
$$\$40,000 \geq P - N_C$$

Clearly $P - N_C$ can't be simultaneously greater and less than \$40,000 and so negotiation costs greater than \$10,000 will preclude the socially optimal outcome (and lead to the more expensive way of cleaning up the pollution problem—that is, Haddonfield paying \$50,000 to filter the incoming water).

The important lesson is that inclusion of transaction costs can change some of our outcomes. If negotiation costs were less than \$10,000, the most efficient solution (Cherry Hill filtering for \$40,000) would occur.

When we considered sophisticated pricing in Chapters 9 and 10, it was necessary to “segment and seal” the markets. The segmenting entails more sophisticated demand analysis than a single price monopoly requires. Such analysis requires more time and money to carry out. First-degree price discrimination (Chapter 9) may require bargaining (haggling) between the buyer and the seller. This requires more salespeople as compared to a single posted price strategy. In addition, the sophisticated pricer must seal the market to prevent resale from a low-price buyer to a high-price buyer. This entails some sort of screening, such as Disney’s technique of requiring visitors to show a driver’s license at the time of entry to their parks. Such screening costs money, for it takes extra time to screen each buyer. In the example of Honest Sanjay in Chapter 9, as long as these extra costs of segmenting and sealing the market are less than \$15.5, Sanjay will practice first-degree price discrimination. Since so many real world firms practice price discrimination, we conclude that the transaction costs of doing so are exceeded by the benefits of practicing sophisticated pricing. One of the advantages of auctions cited in Chapter 13 is the elimination of haggling transaction costs for items when the seller had a difficult time identifying the reservation price of the potential buyers.

Firms have used transaction costs to lock in customers. Consider the long-term cell phone contracts. With a two-year contract with company X, you’re less likely to switch to company Y when you see Y’s new phone and rates, which might appear attractive to a new buyer. That’s because the penalty costs associated with breaking the contract with X exceed the benefits of the new phone/plan with Y. Most consumers, when confronted by a side-by-side comparison of Colgate and Crest toothpastes in the grocery store in which one brand was \$0.50 more expensive, would switch brands in a heartbeat to the cheaper one. We should all remember our lessons about sunk costs when doing our benefit/transaction cost analysis.

Consider the transaction costs your instructor faces when switching textbooks. She is already very familiar with her current adoption. She has tailored her notes and examples to the book. A new text will have different notation and examples. In addition, publishers provide slide decks, teaching guides, test banks, links to current events relevant to the text. All of these services are available with a new text. But adopting a new text means learning how to use these new systems and revising ones notes. These are costs. Are the benefits of another text enough to offset these costs of switching? The same situation occurs when one contemplates switching from one word-processing system to another or from one spreadsheet program to another.

Game theory (Chapter 12) involves strategic thinking. How can I strike a deal that is most advantageous to me? Many large railroads sold off their sparsely used branch lines after railroad deregulation in 1980, preferring to become wholesalers concentrating on their mainline business. Branch lines were expensive for union-

ized large railroads to use because of their low-traffic density and their low speeds (due to poor track condition), which ate up available unionized labor hours. Non-unionized short lines took over many of these routes. Virtually no traffic both originates and terminates on the same short line. Where does the short line originate/terminated traffic go/come from? From the large railroads. And from which main lines? Generally they originate from the original owner of the branch line. Many of the short lines have only one connection. This lowers the short line's bargaining power to deal with the shares of the revenue (called *divisions*) from moving the traffic. Even when branch lines have connections with multiple main lines, contracts with the selling (or leasing) main line require that the short line tender traffic to the original owning railroad. This is a "holdup" effect, an increased cost to the short line of doing business. And it's a very clever strategy by the selling railroad, which now delivers and picks up an aggregate of cars (think wholesale) at the junction of the main line and the short line instead of engaging in the previous very costly retail business of picking up and delivering cars one or two at a time.

Transaction costs are also involved in information asymmetries (Chapters 14, 15, and 16). Individuals and firms unaware of these costs may make decisions differently if they know of these costs. Legislation such as truth in lending lets borrowers know the true cost of borrowing. Warnings that smoking cigarettes may be harmful to your health let you know that you may incur health care costs in addition to the cost of the cigarettes. Your insurance coverage may require you to undergo certain medical tests before you are allowed to pay your premium and be insured and may require certain subsequent expenditure or behavioral modifications if you are to continue to be insured. These are all costs a person needs to consider beyond the insurance premium.

Virtually all of the actions discussed in this book entail transactions costs that enable the transaction to occur (or prevent the transaction from occurring). These costs are above and beyond the obvious costs we have discussed. We have concentrated on the simple concept that there is a "cost" or a "price" to simplify the analysis, but the basic concept is always the same: You do something if the benefits exceed the costs, whereas you don't do something if the reverse is true.

For example, consider our discussion of the full costs of transportation later in this chapter on pages 218 and 219. If we consider the volume or freight moved as a function of the cost of moving freight, we will conclude that the two are inversely related. But what is that cost? It's not just the freight rate charged by the transportation mode. It's also the cost of the time in transit (opportunity cost of inventory in transit—after all, goods are dollar bills in disguise), transit time reliability (which affects safety stock—an amount of inventory held in case delivery is not on time), loss and damage cost, packaging cost, and so on. All freight demand modeling looks at all these costs in total, not just the freight rate charged by the carrier.

NETWORK ECONOMIES

Many industries involve networks. Can you imagine Alexander Graham Bell with his telephone asking Watson “Are you there,” if Watson didn’t have a telephone? Or if American Airlines only flew from its hub at Dallas-Fort Worth to Dallas-Fort Worth? Or if you went on Facebook and found you were the only one there? What if you could only tweet to yourself? The list goes on and on. These industries only have value because multiple people can participate. The larger the network, the more people can participate, and hence the greater their value. Because the secret of the value is in the size of the network, the first entrant in the market can secure a great advantage.

Suppose a world of six people, two of which are the above-mentioned misters Bell and Watson. If just those two have phones, then while they can call each other; the four others can’t be called and none of the other four can call each other. But if Bell can get the other four people to join his phone network, then everyone can talk to everyone else. This means that $6 \times 5 = 30$ possible phone connections can be made. Going from two people with phones to six people with phones leads to a 15 times increase in the amount of potential phone calls. The cost increase to serve the additional customers can be less than 15 times (especially if the people are close together). Thus Mr. Bell can take advantage of economies of scale and scope. He has a fixed investment of running wires from each customer’s home to a phone routing center (that directs the call from person A to person B), but after that investment is made, adding additional customers entails minor adjustments to the routing center and additional wire from the street to the new customer’s house.

Suppose a seventh customer moves into the area and a potential new entrant is considering starting a phone service. The new entrant would have to duplicate Mr. Bell’s investment and then convince potential customers that they had someone they could call. But if Mr. Bell has the first six customers tied up, the new consumer won’t join the new company (there’s no one to call) and would likely join Mr. Bell’s network. Note that Mr. Bell’s network has now become even more attractive to his existing customers because now there are 42 ($= 7 \times 6$) possible connections to be made. Adding the seventh customer (a 16.67% increase) increased potential output by 40%. And these advantages for the incumbent keep rising. Adding one more customer when there are already 100 customers is a 1% increase in customers. But potential connections rise from 100×99 to 100×101 , a 2.02% increase. If n is the current number of customers, an increase of one more customer (a $100/n$ percent increase) will increase the number of potential connections by $200/(n - 1)$ percent. The advantage continues (and only approaches zero as n approaches infinity).

Let’s define a network. To do so, and continue with the logic shown above, we’ll consider a hypothetical example of Federal Express. Suppose that five com-

munities exist that require overnight air service for one unit of product. Four of them (A through D) lie along a square with length 1 and the fifth (E) lies at the intersection of the two main diagonals. The scenario is shown in Figure 6.7.

The distances between each city pair are shown on the figure. The distances from A to D and from B to C are 1.414.

Suppose that aircraft are very expensive and that to travel one distance unit takes one time unit. Each plane can travel 1.6 distance units in the time that it takes to get a good from its origin to its destination absolutely-positively overnight. Also assume that each city wants to ship one unit of product to each other city. Planes can be purchased with any amount of capacity with larger planes being more expensive to run but with costs increasing less proportional to size because pilot/co-pilots are necessary on any aircraft. Operating costs of an aircraft that can carry four units of product are one per distance unit while operating costs on aircraft that can carry one unit of product are 0.75 per distance unit.

One possibility (but not the only one) is to provide point-to-point service for each city pair. If this is done, there are 5×4 or 20 shipments required each night. This would require 20 planes and a total cost shown in Table 6.4. Each plane would operate with capacity of one. The operating cost of running this system is 14.484.

Suppose, alternatively, that point E is treated as a hub airport. All four units shipped from A (destined to B, C, D, and E) are placed on one plane with operating costs of one per distance unit. The same is true for all four shipments from B and from C and from E. Once at E, the E destined shipments are culled out. The A to B shipment is then combined with the C to B shipment and the D to B shipment,

FIGURE 6.7

The Network for a Hypothetical Federal Express

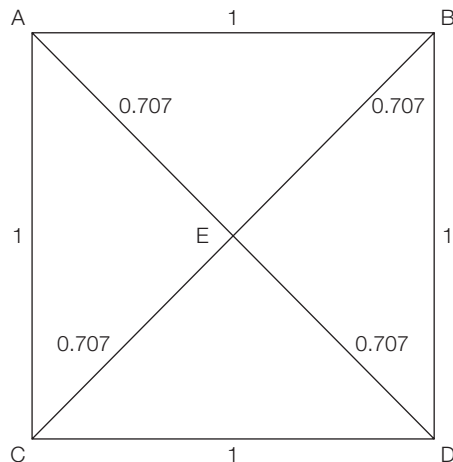


TABLE 6.4

Operating Cost if Each City Pair was Served in Point-to-Point Service

Route	Operating Cost
A to B	$0.75 = 1 \times 0.75$
B to A	0.75
A to C	0.75
C to A	0.75
A to D	$1.0605 = 0.75 \times 1.414$
D to A	1.0605
A to E	$0.53025 = 0.75 \times 0.707$
E to A	0.52025
B to C	1.0605
C to B	1.0605
B to D	0.75
D to B	0.75
B to E	0.53025
E to B	0.53025
C to D	0.75
D to C	0.75
C to E	0.53025
E to C	0.53025
D to E	0.53025
E to D	0.53025
Total	14.484

and the E to B shipment and all four units are shipped to B on the aircraft that came from B. The same is true for all shipments destined to A, C, and D. No plane has exceeded the 1.6 constraint, so long as this sorting at E takes less than $1.6 - 1.414$ in time. So four larger aircraft can now do the job of 20 smaller aircraft (or think that four pilots and co-pilots can now do the job of 20 pilots and co-pilots). Under this scenario, the total costs are shown in Table 6.5.

Using the hub network involves a 61% operating cost saving to provide the same output as the point-to-point service because you can do it with 20% of the aircraft and 20% of the pilots/co-pilots. When the aircraft cost, the operating cost of the aircraft, and the operating cost of the hub are combined, it turns out that it's the cheapest way to move the traffic subject to the absolutely, positively overnight constraint.

TABLE 6.5

Operating Costs if Each City is Served Via a Hub at E

Flight	Cost
A to E	$0.707 = 1 \times 0.707$
E to A	0.707
B to E	0.707
E to B	0.707
C to E	0.707
E to C	0.707
D to E	0.707
E to D	0.707
Total	5.656

MANAGERIAL USE OF BREAK-EVEN ANALYSIS

The concept of breaking even is a cost-based analysis that is both useful and simple. Break-even analysis looks at the relative positioning of costs and revenues; managers use it to estimate how possible pricing changes affect firm performance. Figure 6.8 shows the situation facing Martin divisional managers. They face fixed costs of \$600,000 per month and average variable costs of \$2 per unit. They sell their component for \$3 per unit. Because average variable cost is constant, the cost of an extra unit (marginal cost) is also constant and equal to average variable cost. Given the \$3 per-unit price, the revenue curve is a straight line through the origin. Martin managers create a break-even chart by plotting the firm's total revenue curve with its total cost function. The chart estimates monthly profit for all possible sales. For example, Figure 6.8 shows that if Martin managers have sales of 300,000 units per month, managers realize a loss of \$300,000. The chart also estimates the **break-even point**, which is the output level that must be reached if managers are to avoid losses. This is the intersection of the cost and revenue functions; in Figure 6.8 the break-even point is 600,000 units. A useful way to represent the difference between a product's price and its average variable costs is as the money needed to "cover" the fixed costs. Once managers cover their fixed costs, the difference represents the profit per unit. For the Martin managers, each unit they sell can cover a dollar of fixed costs (\$3 – \$2). Because fixed costs are \$600,000, they need to sell 600,000 units given this \$1 difference. At the 600,000th unit, the firm's profit is \$0; and after 600,000 units are sold, each unit increases profit by \$1.

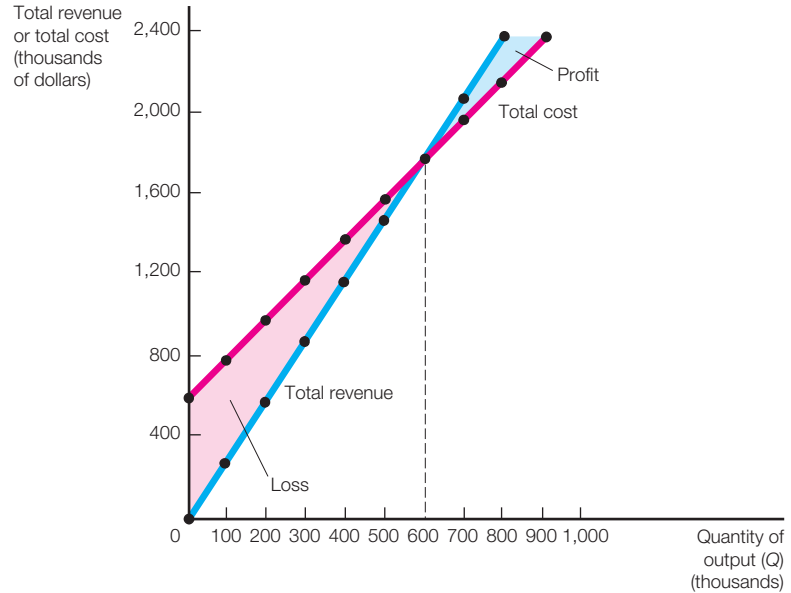
Break-even analysis offers useful estimates of the relationships among sales and costs, receipts, and profit. For example, managers use this analysis to predict how a projected decline in sales will impact profit. Or they use it to estimate how

Break-even point The output level that must be reached if managers are to avoid losses.

FIGURE 6.8

Break-Even Chart: Martin Company

The break-even point—the output level that must be reached if the firm is to avoid losses—is 600,000 units of output per month.

**QUANT OPTION**

Let P be the price of the component, Q the quantity sold, AVC the average variable cost, and TFC the fixed cost. The break-even unit is the output, Q_B , at which $TR = TC$. Because TR is PQ and total cost is $TFC + AVC(Q)$, it follows that

$$\begin{aligned}
 PQ_B &= TFC + AVC \times [Q_B] \\
 (P - AVC)Q_B &= TFC \\
 Q_B &= \frac{TFC}{P - AVC} \quad (6.7)
 \end{aligned}$$

In the case of the Martin managers, $P = \$3$, $AVC = \$2$, and $TFC = \$600,000$. Consequently,

$$Q_B = \frac{\$600,000}{3 - 2} = 600,000$$

a price change might affect profit. Like most models, the analysis has limitations. High variance in prices or difficulties in estimating the cost structure can decrease the accuracy of results.

It is worth noting that although we worked with a linear cost function, this assumption is easily relaxed. Often a curvilinear cost function is used to estimate the cost structure. Many times, for fairly small changes in output, a linear approximation is good enough. This guideline is supported by many empirical studies suggesting that cost functions are often close to linear, as long as managers are not operating at or close to capacity.

PROFIT CONTRIBUTION ANALYSIS

Managers also use break-even analysis to understand the relationship between price and profit; this analysis is known as **profit contribution analysis**. As we have already discussed, profit contribution is the difference between total revenue and

Profit contribution analysis A break-even analysis to understand the relationship between price and profit.

STRATEGY SESSION: Mr. Martin Gets Chewed Out by the Boss

John Martin, an accountant at a small company that manufactures and sells three types of desks, constructed a break-even chart for the company as a whole. He used as a measure of output the total number of desks produced during each year. To estimate the average variable cost of a desk, he took the mean of the average variable costs of the three types of desks. To estimate the price of a desk, he took the mean of the prices of the three types of desks. Using these figures, he constructed a break-even chart (based on linear total cost and total revenue curves) indicating that the company was operating at an output level well above the break-even point and that profit would increase rapidly as output increased.

When Martin presented these findings, the company's president, Susan Rogers, said they were misleading because the analysis lumped together the three types of desks. For one type, the plant was operating at capacity, and marginal cost would increase substantially if output were increased. For another type of desk, it had become increasingly

obvious that the price was too high, and it was about to be reduced. For the third type of desk, only a few were produced, so it was incorrect to weight it as heavily as the other two types in the analysis. Rogers also pointed out that as the firm's output increased, the first and second types accounted for bigger and bigger shares of total output.

How should Martin respond to these comments? (Sometimes it is best just to admit you are wrong and thank the president for the information.) The fact that the product mix changes with increases in output is important, and Martin should have recognized this. It is also misleading to lump all three products together. And contrary to the assumption that the total cost curve is linear, the marginal cost of the first type of desk rises with increases in output. Finally, the price Martin used for the second type of desk is not the relevant one.

Note: This section is based on an actual case, although the numbers and situation have been disguised somewhat.

total variable cost; on a per-unit basis, it is equal to price minus average variable cost. In the case of the Martin Company, price is \$3 and average variable cost is \$2, so the per-unit profit contribution is \$3 − \$2, or \$1.

Suppose managers at Martin want to estimate the sales level that will earn them a profit of \$1 million per month. The required sales equal

$$\begin{aligned} Q &= \frac{\text{Total fixed cost} + \text{Profit target}}{\text{Profit contribution (per unit)}} \\ &= \frac{\$600,000 + \$1,000,000}{\$1} \\ &= 1,600,000 \text{ units} \end{aligned}$$

Or if managers sell only 500,000 units per month, the firm loses \$100,000. The marketing team hope to land an order for 50,000 units of product. How much will this order reduce the firm's loss? To find out, multiply the order size (50,000 units) by the per-unit profit contribution (\$1) to get the increase in profit (or reduction in loss, which is the case here); the result is \$50,000.

SUMMARY

1. Managerial economists define a product's opportunity cost as the value of other products that could have been produced with the money used to produce the product. Hence a product's opportunity cost may differ from its historical cost, which is generally the basis for accounting statements.

2. In the short run it is important to distinguish between a firm's fixed and variable costs. Managers should be able to chart total, average, and marginal costs against output. The resulting cost functions, or cost curves (as they are often called), show how changes in output affect a firm's costs.

3. The long-run average cost function shows the minimum cost per unit of producing a given output level when any desired scale of plant can be built. The long-run average cost function is tangent to each of the short-run average cost functions at the output where the plant corresponding to the short-run average cost function is optimal. The long-run average cost curve is important to managers because it shows the extent to which larger plants have cost advantages over smaller ones.

4. Economies of scope occur when the cost of producing two (or more) products jointly is less than the cost of producing them separately. Such economies may arise because the production facilities used to make one product can also be used to make another product, or by-products resulting from the making of one product can be useful in making other products.

5. Break-even analysis compares total revenue and total cost, graphically or algebraically. A break-even chart combines the total cost function and the total

revenue curve, both of which are generally assumed to be linear, and shows the profit or loss resulting from each sales level. The break-even point is the sales level that must be achieved if the firm is to avoid losses. Managers often find it useful to carry out various types of profit contribution analysis. The profit contribution is the difference between total revenue and total variable cost; on a per-unit basis, it is equal to price minus average variable cost.

PROBLEMS

1. An MIT study has estimated costs for producing steel with three different technologies: (1) coke, blast furnace, basic oxygen furnace, ingots, and finishing mills; (2) coke, blast furnace, basic oxygen furnace, continuous casting, and finishing mills; and (3) steel scrap, electric arc furnace, continuous casting, and finishing mills. Under reasonable assumptions concerning input prices, the estimated average costs per ton are as follows:

Cost Category	Coke, Blast Furnace, Basic Oxygen Furnace, Ingots, Finishing Mills	Coke, Blast Furnace, Basic Oxygen Furnace, Continuous Casting, Finishing Mills	Steel Scrap, Electric Arc Furnace, Continuous Casting, Finishing Mills
Process materials	\$148.34	\$136.19	\$122.78
Energy	21.15	15.98	41.58
Direct labor	83.43	75.09	67.43
Capital	102.06	99.93	54.08
Other	46.74	41.67	24.47
Total	\$401.72	\$368.86	\$310.34

- a. The MIT report concludes that “unless significant changes occur in other technologies, the electric-furnace continuous-casting route will dominate domestic production.” Why?
- b. At the same time, the report notes that the price of scrap (which is used in this route) “could increase as electric furnace production expands because of the increased demand.” Why is this relevant?
- c. The report also concludes that regardless of which of these technologies is used, cost per ton is about 25 to 30% higher if wages are \$26 per hour rather than \$2 per hour. What does this imply about the competitiveness of U.S. steel producers relative to producers in other countries that pay wages far below U.S. levels?
- d. If these cost figures are long-run average costs, under what circumstances would they also equal long-run marginal costs?

2. The Haverford Company is considering three types of plants to make a particular electronic device. Plant A is much more highly automated than plant B, which in turn is more highly automated than plant C. For each type of plant, average variable cost is constant so long as output is less than capacity, which is the maximum output of the plant. The cost structure for each type of plant is as follows:

Average Variable Costs	Plant A	Plant B	Plant C
Labor	\$1.10	\$2.40	\$3.70
Materials	0.90	1.20	1.80
Other	0.50	2.40	2.00
Total	\$2.50	\$6.00	\$7.50
Total fixed costs	\$300,000	\$ 75,000	\$25,000
Annual capacity	200,000	100,000	50,000

- Derive the average costs of producing 100,000, 200,000, 300,000, and 400,000 devices per year with plant A. (For output exceeding the capacity of a single plant, assume that more than one plant of this type is built.)
 - Derive the average costs of producing 100,000, 200,000, 300,000, and 400,000 devices per year with plant B.
 - Derive the average costs of producing 100,000, 200,000, 300,000, and 400,000 devices per year with plant C.
 - Using the results of parts (a) through (c), plot the points on the long-run average cost curve for the production of these electronic devices for outputs of 100,000, 200,000, 300,000 and 400,000 devices per year.
3. The Abner Corporation, a retail seller of television sets, wants to determine how many television sets it must sell to earn a profit of \$10,000 per month. The price of each television set is \$300, and the average variable cost is \$100.
- What is the required sales volume if the Abner Corporation's monthly fixed costs are \$5,000 per month?
 - If the firm sells each television set at a price of \$350 rather than \$300, what is the required sales volume?
 - If the price is \$350, and if average variable cost is \$85 rather than \$100, what is the required sales volume?
4. According to a statistical study, the following relationship exists between an electric power plant's fuel costs (C) and its eight-hour output as a percentage of capacity (Q)

$$C = 16.68 + 0.125Q + 0.00439Q^2$$

- When Q increases from 50 to 51, what is the increase in the cost of fuel for this electric plant?
- Of what use might the result in part (a) be to the plant's managers?

- c. Derive the marginal (fuel) cost curve for this plant, and indicate how it might be used by the plant's managers.
5. The following table pertains to the Lincoln Company. Fill in the blanks:

Output	Total Cost	Total Fixed Cost	Total Variable Cost	Average Fixed Cost	Average Variable Cost
0	50	_____	_____	_____	_____
1	75	_____	_____	_____	_____
2	100	_____	_____	_____	_____
3	120	_____	_____	_____	_____
4	135	_____	_____	_____	_____
5	150	_____	_____	_____	_____
6	190	_____	_____	_____	_____
7	260	_____	_____	_____	_____

6. The Deering Manufacturing Company's short-run average cost function in 2012 was

$$AC = 3 + 4Q$$

where AC is the firm's average cost (in dollars per pound of the product), and Q is its output rate.

- Obtain an equation for the firm's short-run total cost function.
 - Does the firm have any fixed costs? Explain.
 - If the price of the Deering Manufacturing Company's product (per pound) is \$3, is the firm making profit or loss? Explain.
 - Derive an equation for the firm's marginal cost function.
7. The president of the Tacke Corporation believes that statistical research by his staff shows that the firm's long-run total cost curve can be represented as

$$TC = \alpha_0 Q^{\alpha_1} P_L^{\alpha_2} P_K^{\alpha_3}$$

where TC is the firm's total cost, Q is its output, P_L is the price of labor, and P_K is the price of capital.

- Tacke's president says that α_1 measures the elasticity of cost with respect to output—that is, the percentage change in total cost resulting from a 1% change in output. Is he correct? Why or why not?
- He also says that if $\alpha_1 < 1$, economies of scale are indicated, whereas if $\alpha_1 > 1$, diseconomies of scale are indicated. Is he correct? Why or why not?
- According to Tacke's president, the value of α_3 can be estimated by regressing $\log(TC/P_K)$ on $\log Q$ and $\log(P_L/P_K)$. Is he correct? Why or why not?

8. Engineers sometimes rely on the “0.6 rule,” which states that the increase in cost is given by the increase in capacity raised to the 0.6 power; that is,

$$C_2 = C_1(X_2/X_1)^{0.6}$$

where C_1 and C_2 are the costs of two pieces of equipment, and X_1 and X_2 are their respective capacities.

- Does the 0.6 rule suggest economies of scale?
 - Some experts have stated that in the chemical and metal industries, the 0.6 rule can be applied to entire plants rather than individual pieces of equipment. If so, will the long-run average cost curve in these industries tend to be negatively sloped?
 - Can you think of a way to test whether this rule is correct?
9. The Dijon Company’s total variable cost function is

$$TVC = 50Q - 10Q^2 + Q^3$$

where Q is the number of units of output produced.

- What is the output level where marginal cost is a minimum?
 - What is the output level where average variable cost is a minimum?
 - What is the value of average variable cost and marginal cost at the output specified in the answer to part (b)?
10. The Berwyn Company is considering the addition of a new product to its product line. The firm has plenty of excess manufacturing capacity to produce the new product, and its total fixed costs would be unaffected if the new product were added to its line. Nonetheless, the firm’s accountants decide that a reasonable share of the firm’s present fixed costs should be allocated to the new product. Specifically, they decide that a \$300,000 fixed charge will be absorbed by the new product. The variable cost per unit of making and selling the new product is \$14, which is composed of the following:

Direct labor	\$8.20
Direct materials	1.90
Other	3.90
Total	<u>\$14.00</u>

- Should the Berwyn Company add the new product to its line if it can sell about 10,000 units of this product at a price of \$25?
- Should it add the new product if it can sell about 10,000 units at a price of \$20?
- Should it add the new product if it can sell about 10,000 units at a price of \$15?

- d. What is the minimum price for the new product that will make it worthwhile for Berwyn to add the new product to its line?
11. The Jolson Corporation produces 1,000 wood cabinets and 500 wood desks per year, the total cost being \$30,000. If the firm produced 1,000 wood cabinets only, the cost would be \$23,000. If the firm produced 500 wood desks only, the cost would be \$11,000.
- Calculate the degree of economies of scope.
 - Why do economies of scope exist?
12. The Smith Company made and sold 10,000 metal tables last year. When output was between 5,000 and 10,000 tables, its average variable cost was \$24. In this output range, each table contributed 60% of its revenue to fixed costs and profit.
- What was the price per table?
 - If the Smith Company increases its price by 10%, how many tables will it have to sell next year to obtain the same profit as last year?
 - If the Smith Company increases its price by 10%, and if its average variable cost increases by 8% as a result of wage increases, how many tables will it have to sell next year to obtain the same profit as last year?

EXCEL EXERCISE: PRODUCTION AND COST

Suppose that your industrial engineers presented you with the following technically efficient ways of combining labor (L) with the two units of capital (K) you possess.

Capital	Labor	Quantity
2	0.00	0
2	0.25	1
2	0.60	2
2	1.05	3
2	1.60	4
2	2.25	5
2	3.00	6
2	3.85	7
2	4.80	8

You also know that the cost of a unit of capital (r) is 2.5 and that the cost of a unit of labor (w) is 10. Is this enough information to derive the firm's total cost function (and other associated costs of the firm)? The answer is yes; let's see how the spreadsheet eases that calculation.

Open up your spreadsheet. Enter 2 in cell A1. Then click the lower right hand corner of cell A1 and drag down until you reach cell A9, then release. This should leave with a column of nine 2s.

Then enter 0 in cell B1, 0.25 in cell B2, 0.6 in cell B3, and so on, until you've entered 4.8 in cell B9.

Then enter 0 in cell C1, 1 in cell C2, and so on, until you've entered 8 in cell C9.

You have now replicated the table in the spreadsheet.

Then enter $=C2/B2$ in cell D2, $=C3/B3$ in cell D3, and so on. You can use the click-and-drag method to fill all the cells from D3 to D9. You have calculated the average product of labor—that is, $AP_L = Q/L$, the famous labor productivity measure.

Then enter $=(C2-C1)/(B2-B1)$ in cell E2 and $=(C3-C2)/(B3-B2)$ in cell E3, and so on. You can use the click-and-drag method to fill all the cells from E3 to E9. You have calculated the marginal product of labor—that is, $MP_L = \Delta Q/\Delta L$.

Then enter 10 in cell F1 and drag and click until you reach cell F9. That should give you a column of 10s, which is the wage rate the firm faces for each unit of labor it hires (and as can be seen from column 2, it can hire labor in fractional units—think part-time workers).

Then enter $=F2/E2$ in cell G2, $=F3/E3$ in cell G3, and so on via the click-and-drag method, until you reach cell G9. You now have the firm's marginal cost in column G. Recall from the text that $MC = w/MP_L$.

Then enter $=F2/D2$ in cell H2, $=F3/D3$ in cell H3, and so on via the click-and-drag method, until you reach cell H9. You now have the firm's average variable cost in column H. Recall from the text that $AVC = w/AP_L$.

Then enter 2.5 in cell I1 and drag and click till cell I9. You should have a column of 2.5s which is the cost of a unit of capital.

Then enter $=A1*I1$ in cell J1, and $=A2*I2$ in cell J2, and so on via the click-and-drag method. This should give you a column of 5s, the firm's fixed cost ($FC = rK$).

Then enter $=B1*F1$ in cell K1 and $=B2*F2$ in cell K2, and so on via click-and-drag. This will give you the firm's variable cost ($VC = wL$).

Then enter $=J1+K1$ in cell L1, and $=J2+K2$ in cell L2, and so on via click-and-drag. This will give you the firm's total cost ($TC = FC + VC$).

You might even be able to discern a relationship between variable cost and quantity. Note that every number in column K (using the Q s in column C) can be calculated from the formula $VC = 2Q + 0.5Q^2$. Thus, the firm's total cost function is $TC = 5 + 2Q + 0.5Q^2$.

But you don't have to know that relationship because you were able to enumerate the firm's costs by using the production function given to you by your industrial engineers in the table above along with the unit prices of your inputs, r and w .

APPENDIX A: BREAK-EVEN ANALYSIS AND OPERATING LEVERAGE

Managers must continually compare alternative systems of production. Should one type of plant be replaced by another? How does your plant stack up against your competitor's? Break-even is used to help make such comparisons more effective. In this appendix we show managers how to analyze changes in total cost and profit, depending on how automated or mechanized a plant may be. This is an important topic because top-level managers often have to make such comparisons.

At the outset it is essential to recognize that some plants, because they are more mechanized than others, have relatively high fixed costs but relatively low average variable costs. Consider firms I, II, and III in Figure 6.9. Firm I's plant has fixed costs of \$100,000 per month, which are much higher than those of the plants operated by firm II or III; however, its average variable cost of \$2 is much lower than that of firm II or III. Essentially firm I has substituted capital for labor and materials. Managers have built a highly automated plant with high fixed costs but low average variable cost.

At the opposite extreme, managers of firm III have built a plant with low fixed costs but high average variable cost. Because they have not invested a great deal in plant and equipment, total fixed costs are only \$25,000 per month, which is much less than that for firm I or II. However, because of the relatively low level of automation at its plant, firm III's average variable cost is \$4—considerably higher than at the other two firms. Relative to firm I, firm III uses more labor and materials and less capital.

Firm II's plant occupies a middle position (between firms I and III) in this regard. Its total fixed cost of \$60,000 is less than firm I's but more than firm III's, and its average variable cost of \$3 is greater than firm I's but less than firm III's. It has not automated its plant to the extent that firm I has, but it has automated more than firm III.

In comparing these plants, an important issue to consider is the degree of operating leverage, which is defined as the percentage change in profit resulting from a 1% change in the number of units of product sold. Specifically,

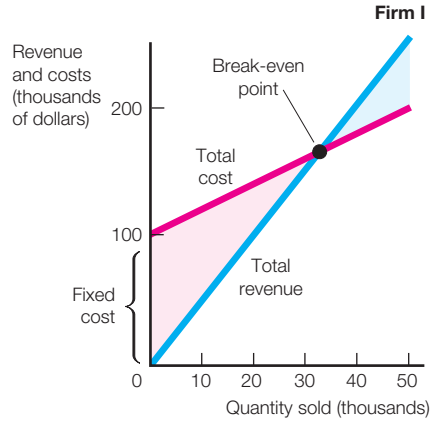
$$\begin{aligned}
 \text{Degree of operating leverage} &= \frac{\text{Percentage change in profit}}{\text{Percentage change in quantity sold}} \\
 &= \frac{\Delta \pi / \pi}{\Delta Q / Q} \\
 &= \frac{\Delta \pi}{\Delta Q} \left(\frac{Q}{\pi} \right) \quad \text{or} \quad \frac{d\pi}{dQ} \left(\frac{Q}{\pi} \right) \quad (6.8)
 \end{aligned}$$

where π is the firm's profit, and Q is the quantity sold.

FIGURE 6.9

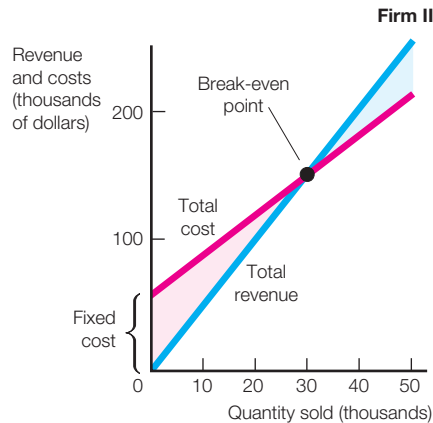
Break-Even Analysis and Operating Leverage

Firm I has relatively high fixed costs and low variable costs; firm III has relatively low fixed costs and high variable costs; and firm II is in the middle.



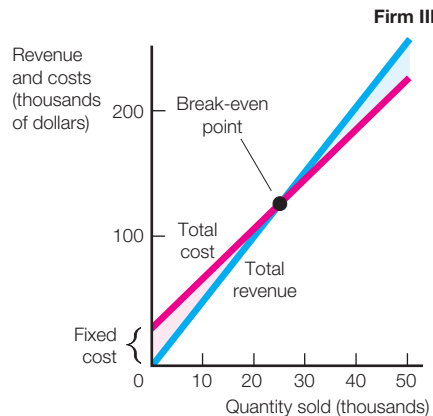
Total fixed cost = \$100,000
 Average variable cost = \$2
 Selling price = \$5

Quantity sold	Total revenue	Total cost	Total profit
10,000	\$ 50,000	\$120,000	\$-70,000
20,000	100,000	140,000	-40,000
30,000	150,000	160,000	-10,000
40,000	200,000	180,000	20,000
50,000	250,000	200,000	50,000



Total fixed cost = \$60,000
 Average variable cost = \$3
 Selling price = \$5

Quantity sold	Total revenue	Total cost	Total profit
10,000	\$ 50,000	\$ 90,000	\$-40,000
20,000	100,000	120,000	-20,000
30,000	150,000	150,000	0
40,000	200,000	180,000	20,000
50,000	250,000	210,000	40,000



Total fixed cost = \$25,000
 Average variable cost = \$4
 Selling price = \$5

Quantity sold	Total revenue	Total cost	Total profit
10,000	\$ 50,000	\$ 65,000	\$-15,000
20,000	100,000	105,000	-5,000
30,000	150,000	145,000	5,000
40,000	200,000	185,000	15,000
50,000	250,000	225,000	25,000

The degree of operating leverage, because it measures how a given change in sales volume affects profit, is of great importance. If firm I is selling 40,000 units per month, and if we let $\Delta Q = 10,000$ units, the degree of operating leverage equals

$$\frac{\Delta \pi}{\Delta Q} \left(\frac{Q}{\pi} \right) = \frac{\$50,000 - \$20,000}{10,000} \left(\frac{40,000}{\$20,000} \right) = 6$$

because Figure 6.9 shows that if $\Delta Q = 10,000$ units, $\Delta \pi = \$50,000 - \$20,000$. (Why? Because if Q changes from 40,000 to 50,000 units, π changes from \$20,000 to \$50,000.) Thus a 1% increase in quantity sold gives a 6% increase in profit.

If both the total revenue curve and the total cost function are linear, as in Figure 6.9, a simple way to calculate the degree of operating leverage when output equals Q is to use the following formula

$$\text{Degree of operating leverage} = \frac{Q(P - AVC)}{Q(P - AVC) - TFC} \quad (6.9)$$

where P equals selling price, AVC equals average variable cost, and TFC equals total fixed cost. It can be shown that if both the total revenue curve and the total cost function are linear, equation (6.9) yields the same result as equation (6.8). Thus for firm I, if $Q = 40,000$, equation (6.9) says that the degree of operating leverage equals

$$\begin{aligned} \frac{Q(P - AVC)}{Q(P - AVC) - TFC} &= \frac{\$40,000(\$5 - \$2)}{40,000(\$5 - \$2) - \$100,000} \\ &= \frac{\$120,000}{\$120,000 - \$100,000} = 6 \end{aligned}$$

because P equals \$5, AVC equals \$2, and TFC equals \$100,000. The result is the same as in the previous paragraph. (In both cases it is 6.)

It is interesting and important to compare the degree of operating leverage of the three firms; this comparison reveals a great deal about how the operation of each plant differs from that of the other. If $Q = 40,000$, the degree of operating leverage for firm II equals

$$\frac{Q(P - AVC)}{Q(P - AVC) - TFC} = \frac{\$40,000(\$5 - \$3)}{40,000(\$5 - \$3) - \$60,000} = 4$$

For firm III, it equals

$$\frac{Q(P - AVC)}{Q(P - AVC) - TFC} = \frac{\$40,000(\$5 - \$4)}{40,000(\$5 - \$4) - \$25,000} = 2.67$$

STRATEGY SESSION: Water, Water, Not Everywhere

In an economy that still has a number of command-and-control aspects, one way to handle a water shortage in China would be to ration supply by decreeing what amount one could use. But markets are growing in China, so another solution is to raise the price, and that's what just about every major city has done in the last decade; they need to conserve water because of the great demand relative to supply, but this requires high prices, and the constituency is accustomed to low water prices. Prices are scheduled to rise by 40 to 48% in Luoyang. Shanghai raised its prices by 25% in June 2009 and planned a 22% increase in November 2010. Zhengzhou raised its water fees by 25% in April 2009 and says that fees will have to change more rapidly in the future. In Nanjing, rates rose by 12% in April 2009. Several other major cities have also recently raised their rates. Even in cities that didn't recently raise rates, such as Beijing (which almost doubled its water rates five years ago), citizens think that water is already expensive and don't want to see increases.

Chinese officials now believe that low prices are a major part of China's water shortage problem. The low prices give little incentive for residences and business to use water prudently. At one time, it was estimated that 20% of supply was lost through leaky pipes.

To see how China's water prices compare with the rest of the world, Deutsche Bank estimated the price of a cubic meter of water, expressed in U.S. dollars, in major developed countries throughout the world. The table below shows the results.

Germany	\$3.01
United Kingdom	\$2.37
France	\$2.00
Australia	\$1.82
Italy	\$1.58
Canada	\$1.02
South Africa	\$1.02
United States	\$0.74
Brazil	\$0.65
China	\$0.31

Thus a 1% increase in sales volume results in a 6% increase in profit at firm I, a 4% increase in profit at firm II, and a 2.67% increase in profit at firm III. Clearly firm I's profit are much more sensitive to changes in sales volume than are firm III's profit; firm II is in the middle in this regard.

APPENDIX B: MEASUREMENT OF SHORT-RUN COST FUNCTIONS: THE CHOICE OF A MATHEMATICAL FORM

Smart managers understand the need to estimate cost functions for their informational value. In business these are often called cost curves. One step in estimating a cost curve is to choose the mathematical relationship between output and cost. As a first approximation, managers often assume that short-run total cost is

The cost of water in Germany is 10 times that in China. Of course these prices reflect differing supply-and-demand situations in each country and different pricing levels for all goods and services. Nevertheless, the consensus is that water is priced too low in China. Per-capita water availability is one quarter of the world average. A study by the World Bank estimates that water shortages cost China 1.3% of its yearly economic output. A lot of the water in China is polluted and that further cuts economic output by 1%.^a

There's a lesson in supply and demand here. If prices rose, people would reduce waste, and the allocation of water across users would improve (the allocatable efficiency of the price system) on the demand side. Higher prices would also encourage investment in projects, such as desalination, that would increase supply. In fact, "if China moved more aggressively to price water in a manner that reflected supply and demand," says Peter Orszag, vice chairman of global banking at Citicorp and former head of the federal Office of Management and Budget (OMB), "it could teach the U.S. a lesson in using market economics

to address environmental issues." Orszag notes that water prices in China are much too low to ensure that water is used efficiently enough to sustain the supply. Water has never been priced efficiently in the United States, claims a Citicorp analyst (note our low cost in the table above), it is heavily subsidized. Orszag states that "market forces can work wonders for the environment, but only if we have the political courage to create them."^b

^a Andrew Batson, "China Cities Raise Water Prices in Bid to Conserve," *The Wall Street Journal*, July 31, 2009, at <http://online.wsj.article/SB124897577003694405.html>.

^b Peter Orszag, "Why We Care about the Price of Water in China: Peter Orszag," *Bloomberg News*, July 6, 2011, at www.bloomberg.com/2011-07-06/why-we-care-about-the-price-of-water-in-china-Peter-orszag.html.

a linear function of output, which means marginal cost tends to be constant in the relevant output range (see Figure 6.10). In fact, this simple linear approximation often fits the data for particular firms and plants quite well in the short run. However, managers need to note that although marginal costs may vary little over a wide range of output, it is inconceivable that they do not eventually rise with increases in output. Therefore, a linear function is likely to be appropriate only for a restricted range of output.

It is also possible to assume total cost is a quadratic or cubic function of output. If the quadratic form is chosen, marginal cost increases with output, as shown in Figure 6.11. If the cubic form is chosen (and c is large enough), marginal cost first decreases then increases with output, as shown in Figure 6.12. Whether these forms are better than the linear form depends on whether they fit the data better. In many cases, they fit the data slightly better than the linear form.

FIGURE 6.10

Average Cost and Marginal Cost: Linear Total Cost Function

Marginal cost is constant.

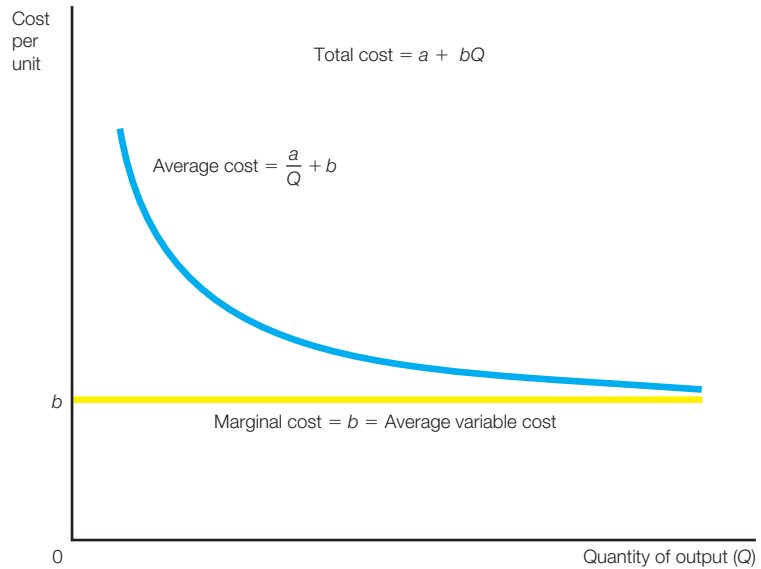


FIGURE 6.11

Average Cost and Marginal Cost: Quadratic Total Cost Function

Marginal cost increases as output rises.

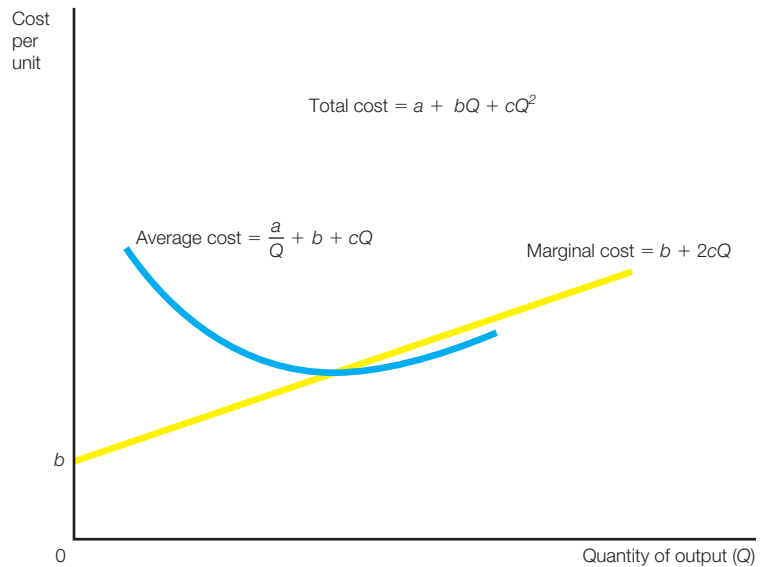
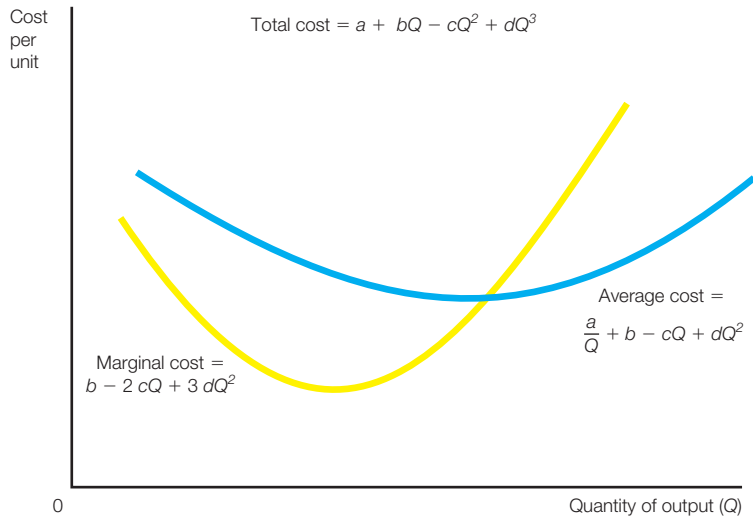


FIGURE 6.12

Average Cost and Marginal Cost: Cubic Total Cost Function

Marginal cost first falls then rises as output increases.

**Key Steps in the Estimation Process**

When managers have chosen a mathematical form and decided on their data set, the following six items should be given careful thought:

1. *Definition of cost:* As we said at the beginning of this chapter, the relevant concept of cost for managers is opportunity cost, not cost based on accounting data. We must be careful to ensure that the accounting data—or engineering data, for that matter—on which an estimated cost function is based are reasonably indicative of opportunity costs. If not, adjust the data. For example, suppose historical data regarding a firm's depreciation costs are based on tax laws rather than on opportunity costs of the relevant equipment. Managers need to revise the cost data to better reflect both opportunity costs and tax conventions.

2. *Correction for price level changes:* When managers use time series data to estimate cost functions, it is important that changes over time in the input prices be recognized and measured. What managers need is a cost function based on next year's input prices if next year is the period to which the analysis pertains. Because historical data are based on input prices at various times in the past, we need a price index to allow us to adjust our historical cost data for changes in various input prices. Moreover, because various inputs may experience quite different rates of inflation, managers should construct a separate price index for

STRATEGY SESSION: The Value of Time and the Full Price of Transportation

Transportation economists have written for years about the full cost of using a transportation mode. It's not just the fare or the gas and parking costs but also the time spent on accessing and egressing the mode; the costs associated with unreliable transport time; safety costs; the ability to pursue other activities while being transported; and so on. Generally the most important cost component (at least for higher-income people) is travel time.

We can get a feel for this by observing what happens in markets where consumers have options. Take State Route 91 in California, for example. Toll lanes were built in the median of an existing freeway. The freeway is heavily congested and often takes two hours to drive in peak periods. In the toll lanes, however, the same trip takes a half hour. The top cost? \$9.45 during the 4 pm to 5 pm hour on Thursdays going eastbound. So what's saving an hour and a half worth? At least \$6.30/hour (that is, $9.45/1.5$), or a rational driver wouldn't use such lanes. The lanes are continuously priced so that a "free flow" (a flow at the posted speed limit) is maintained. This is supply and demand at work. If too many cars want to use the toll lanes so that the posted speed can't be maintained, the toll is increased (and users can choose whether or not to enter the lanes if they wish). Likewise, if the toll lanes can tolerate more cars and still maintain the posted speed limit, the tolls are lowered.

If \$9.45 seems steep, consider that such a policy encourages car pools. Cars with three or more occupants can use the toll lanes for 50% of the toll during evening rush hour and a minimum of \$1.30 at all other times. Cars with multiple occupants can split the toll, mitigating some of its impact. California also is pricing Interstate 15 in San Diego. The states

of Washington, Virginia, and Texas are now considering toll roads. The Bush administration proposed congestion pricing in the 2007 State of the Union message. Singapore has been using congestion pricing for over 36 years, and the system was adopted over eight years ago in London and more recently in Stockholm. Mayor Bloomberg of New York City has proposed an \$8 prime time weekday toll for Manhattan south of 86th Street.

But offers to save time are not limited to highways. You've perhaps experienced a long delay in an airport security line. Would you like to bypass such lines? Even if the line isn't long, you have to report to the airport early because the line might be long. In such cases you spend long times in the secure area of the airport—time you probably would rather spend elsewhere.

A firm called Verified Identity Pass will allow (for a fee of \$99.95 per year) British Airways travelers to avoid the typical security wait. Certain airports (Orlando, Indianapolis, Cincinnati, and San Jose) are interested in establishing lanes where passengers have been precleared and only have to verify their identity by a fingerprint or an iris scan. If you were a businessperson flying the New York–London route 10 times a year, you'd pay perhaps \$10 to avoid a half-hour delay each time you flew. A weekly commuter on the same route would pay only \$2 to save that half hour.

The time spent in security, the airlines' cutting service to small airports, and the fact that virtually all service between small airports is through a hub airport (with the resulting waiting time for connecting flights) is spawning another industry: an air taxi service. With the production of very light jets at the

relatively low price of \$1.8–\$2.4 million, which can fly up to 1,200 miles nonstop and can land at over 5,000 U.S. general aviation airports (there are general aviation airports almost everywhere), several entrepreneurs are contemplating entering the air taxi service market.

Why do they believe this will be lucrative? Because of the full costs of travel via conventional carriers to and from as well as between small markets. Such a service will allow customers to fly directly from origin to destination, transforming many trips to home and back in a day rather than the current two-three day trips with their accompanying meals and hotel accommodations.

Such taxis will be easy to catch at the general aviation airports in the suburbs of big cities; but taxis may be unwilling to fly to locations where generating return fares will not be easy.

Several airlines once flew the lucrative New York–London route offering all business-class service. They priced this service in the \$1,975 range compared with the business-class services offered by the legacy carriers, such as British Airways, Virgin Atlantic, and American Airlines, which charged around \$5,925. While the legacy carriers go to London Heathrow, the new carriers went to Stansted (farther from central London). In addition, connections at Stansted are mostly with discount carriers (like Ryanair); so after transatlantic luxury, a beyond-London traveler would face minimal service.

But why wouldn't a London-destined business traveler want to save \$3,950 on her fare? It's the full pricing again. British Air runs 10 round-trips per day, and Virgin offers five round-trips per day, whereas the all-business-class airlines usually offered one. Many

businesspeople value the ability to go when they want to; this is called schedule delay and is measured by the difference between when the customer wants to go and when the carrier is scheduled to leave. And don't forget the legacy carriers' loyalty awards and frequent-flier miles. One can go anywhere in the world on British Air. Finally, because this service caters to business travelers, they aren't paying the bill (the company is); and a business expense is tax deductible, so the cost difference isn't as great as it initially appears. Alas, focusing on business-class-only customers does have a downside. During the economic crisis of 2008–2012, most business-class-only airlines went bankrupt as business travel plummeted.

Are there hidden full costs for nontransportation goods and services? Because most goods take time to consume and may require other expenditures (say, user assembly time and electricity and repair costs), all rational decision makers should consider the full costs of each product and service.

Sources: Timothy Egan, "Paying on the Highway to Get Out of First Gear," *New York Times*, April 28, 2005, at www.nytimes.com/2005/04/28/natural/28toll.html; Patrick McGeehan, "For a Price, a Faster Way through Local Airports," *New York Times*, December 5, 2006, at www.nytimes.com/2006/12/05/nyregion/05screen.html; Joe Sharkey, "Standing on a Runway Hailing an Air Taxi," *New York Times*, February 28, 2006, at www.nytimes.com/2006/02/28/business/28road.html; Joe Sharkey, "TransAtlantic Luxury for Less," *New York Times*, February 21, 2006, at www.nytimes.com/2006/02/21/business/21compete.html; and "Get Moving on Traffic Relief," *New York Times*, May 25, 2007, at www.nytimes.com/2007/05/25opinion/25fri2.html.

each major input. Using these price indexes, managers can convert the available historical cost data into cost data reflecting next year's input prices, not those of the past.

3. *Relating cost to output*: For an estimated cost function to be reasonably accurate, it is important that cost data distinguish properly between costs that vary with output and those that do not. For many types of equipment, as well as other assets, depreciation depends on both the passage of time and the extent of use, with the result that it is difficult or impossible to determine solely from accounting data how much the depreciation cost varies with output alone. Also, some costs do not vary with output so long as output does not exceed a critical level. Above this critical level, these costs may increase considerably. For example, up to some output level, managers need just one machine tool of a particular type; but beyond that output level they need to obtain an additional machine tool.

4. *Matching time periods*: Major errors sometimes occur because cost data do not pertain to the same time periods as output data. To see what mayhem this can cause, suppose we were to plot a firm's 2012 cost against its 2011 output. Would the resulting chart be a good estimate of the firm's cost function? Of course not. Instead managers need to relate a firm's costs in a particular period to its output in that same period. Managers need to modify this guideline in cases in which some costs of producing output in one period do not arise until subsequent periods. These delayed costs must be recognized, measured, and charged against the period in which the output occurred. For example, the costs of maintenance and repairs, when they are delayed, should be treated in this way.

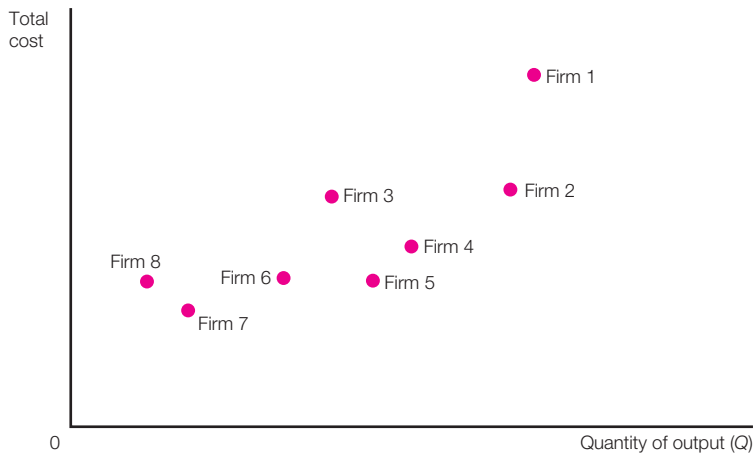
5. *Controlling product, technology, and plant*: Managers need to estimate a firm's cost function on the basis of a fixed definition of the firm's product, as well as on a fixed level of technology and (for short-run cost functions) a fixed scale of plant. This means that managers should be careful to ensure that the firm's product mix does not significantly change over time. Also, the observations used in the analysis should not cover so long a period that they pertain to different levels of technology (or different scales of plant).

6. *Length of period and sample size*: Although managers should prefer a larger sample size, they cannot extend the data gathering phase too long. No simple rule can specify the best length of time. In deciding how long to wait, managers need to consider issues like the level of technology change, seasonal effects, and product changes.

Nature and Limitations of Available Data

Having chosen a mathematical form, managers must select the type of data to use in estimating a cost function. One possibility is to use time series data. Another possibility is to use cross-section data and relate the total costs of a variety of firms (during the same period) to their output levels. Figure 6.13 plots the 2012 output of eight firms in a given industry against their 2012 total costs. Here, too, regres-

FIGURE 6.13

Relationship between Total Cost and Output, Cross Section

sion analysis can be used to estimate the relationship. A third possibility is to use engineering data to construct cost functions.

Regardless of which types of data are used, there are a number of important issues in estimating cost functions. Accounting data, which are generally the only cost data available, suffer from a number of possible deficiencies. Accountants may use arbitrary allocations of overhead and joint costs. The depreciation of an asset is determined largely by tax laws rather than economic criteria. Many inputs are valued at historical cost and do not include opportunity cost.

Engineering data also may contain possible issues. An inherent arbitrariness is involved in allocating costs jointly when producing more than one product in multiproduct firms.

PART 4

MARKET STRUCTURE AND SIMPLE PRICING STRATEGIES

CHAPTER 7

LEARNING OBJECTIVES

Market Structure

Market Price in Perfect
Competition

Shifts in Supply and Demand
Curves

The Output Decision of a
Perfectly Competitive Firm

Setting the Marginal Cost Equal
to the Price

Another Way of Viewing the
Price Equals Marginal Cost
Profit-Maximizing Rule

Producer Surplus in the
Short Run

Long-Run Equilibrium of
the Firm

The Long-Run Adjustment
Process: A Constant-Cost
Industry

The Long-Run Adjustment
Process: An Increasing-Cost
Industry

How a Perfectly Competitive
Economy Allocates Resources

Summary

Problems

Excel Exercise: Perfect
Competition

PERFECT COMPETITION

We are at a natural point to review our path. We first looked at behavior in simple markets and the laws of supply and demand that govern them. We next focused on the supply side of the market, looking at production and cost structures. Our goal was to provide tools and understanding to help managers improve their operating efficiency. Now we want to switch to the demand side and examine managerial behavior and pricing decisions. Only after managers understand both sides of the market can they consider maximizing profit.

Our decision model, like most rational models, assumes managers want to maximize firm value. In the strictest sense this means maximizing profit, though we believe value is more than a simple financial measure. Managers can bring value to their firms by thinking about long-term reputational effects. All decisions have short- and long-term consequences; the foresight to maximize firm value is the benchmark by which managerial ability is judged.

Managers live within a constrained world. They make decisions subject to the limits imposed by technology, resources, economics, politics, and avarice. A dominant constraint in pricing decisions is the market structure. Market structure is important because it largely determines the potential pricing power of managers. We say potential because many managers fail to price optimally. We discuss this further when we introduce sophisticated pricing in Chapters 9 and 10.

Managers should classify markets based on their degree of pricing power. At one end is a perfectly competitive market in which managers have no market

power. The other end is anchored by monopoly markets, where managers face no competition and possess plenty of market power.

We begin by examining pricing behavior in perfectly competitive markets. From a strategic view, these markets are less interesting because individual managers have no effect on price. Instead they are ruled by the “invisible hand” described by Adam Smith. Managers are price takers: They accept the decisions of the aggregate market. However, perfectly competitive markets serve as good benchmarks to evaluate any value created by managers. They also lay the groundwork for our journey through more strategically interesting markets where managers must consider and anticipate the actions of others (consumers and rivals) in setting prices.

In many markets managers operate with no market power. For example, as a small participant in a large market where the market supply and demand curves determine the price, the manager maximizes the firm’s profit given the market-determined price. Other examples occur when the government sets the price (either via price controls or by virtue of being in a regulated industry) or where the firm is a follower firm in a market where a price leader sets the market price. Managers with no market power have no control over price; they are at the whim of the market.

If you find yourself in such a market, here is what you should expect. You will still face the supply-side challenges of efficient production and cost control. This is true for all managers at all times. From the demand side, managers must choose the profit-maximizing output when the price is given. In all other markets, managers can vary both output and price.

In perfectly competitive markets, managers cannot overrule the price set by the interaction of the aggregate market demand and supply curves. An individual manager cannot influence the market price. It is often suggested that the farmer is in this position. While not controlling price, the farmer does decide on quantity—that is, how much of each crop to plant. This decision must be made months before the crop comes to market. To minimize risk, a farmer may sell a crop for future delivery, but most do not. After the crop is planted, the price may change as conditions influencing demand (tastes, income) and supply (weather, crop disease) change. The farmer makes a second quantity decision at harvest time: How much of the crop to harvest. Farmers may also stop growing a crop before it matures, replot, and plant another crop because of changes in market prices. What is common in all these cases is that the farmer makes quantity decisions only, and they are based on a price over which the farmer has no control.

MARKET STRUCTURE

This situation of a price-taking producer is one of four general categories of market structure we investigate. We preview all four categories in this section, then spend the rest of this chapter discussing the perfect competition (price taker)

Perfect competition When there are many firms that are small relative to the entire market and produce similar products.

Monopolistic competition When there are many firms and consumers, just as in perfect competition; however, each firm produces a product that is slightly different from the products produced by the other firms.

Monopoly Markets with a single seller.

Oligopoly Markets with a few sellers.

Barriers to entry Barriers that determine how easily firms can enter an industry, depending on the market structure.

category. As we pointed out in Chapter 1, a market consists of a group of firms and individuals that buy or sell some good or service. Economists have generally found it useful to classify markets into four broad types: perfect competition, monopoly, monopolistic competition, and oligopoly. In **perfect competition** and **monopolistic competition** there are many sellers, each of which produces only a small part of the industry's output. **Monopoly** markets, on the other hand, consist of only a single seller. **Oligopoly** is an intermediate case where there are a few sellers; this is the most prevalent category in present-day business. Hence American Water (which serves 16.2 million customers in 32 states and Ontario), if it is the only supplier of water in a particular market, is a monopoly. And because there are only a few automobile manufacturers, the market for automobiles is an oligopoly.

Market structures vary substantially in the extent to which managers can control price. Managers in perfectly competitive markets have no control over price. For example, a farmer producing corn has no control over the price of corn. On the other hand, a monopolist is likely to have considerable control over price. In the absence of public regulation, American Water would have considerable control over the price of water in the locations it serves. A manager operating under monopolistic competition or oligopoly is likely to have less control over price than a monopolist and more control over price than a manager in a perfectly competitive market.

These market structures also vary in the extent to which the firms in an industry produce standardized (that is, identical) products. Firms in a perfectly competitive market all produce identical products. One farmer's wheat is essentially the same as another farmer's. In a monopolistically competitive industry like shirt manufacturing, firms produce somewhat different products. One firm's shirts differ in style and quality from another firm's shirts. In an oligopoly, firms sometimes, but not always, produce identical products; for example, in steel and aluminum they do, whereas in cars, they do not. And in a monopoly there can be no difference among firms' products because the industry contains only one firm.

How easily firms can enter an industry differs from one market structure to another. In perfect competition **barriers to entry** are low. For example, only a small investment is required to enter many parts of agriculture. Similarly, there are low barriers to entry in monopolistic competition. But oligopolies such as automobile manufacturing and oil refining tend to feature considerable barriers to entry: It is very expensive to build an automobile plant or an oil refinery. In a monopoly entry is blocked; if entry occurs, the monopoly no longer exists.

Market structures also differ in the extent to which managers compete on the basis of advertising, public relations, and different product characteristics, rather than price. In perfect competition there is no nonprice competition. (If every farmer produces identical corn and has to accept the market price, why devote funds to advertising?) In monopolistic competition considerable emphasis is placed on managers using nonprice competition. Much of this nonprice competition centers around the ability of managers to differentiate their products; this

TABLE 7.1

Characteristics of Perfect Competition, Monopolistic Competition, Oligopoly, and Monopoly

Market Structure	Examples	Number of Producers	Type of Product	Power of Firm over Price	Barriers to Entry	Nonprice Competition
Perfect competition	Some sectors of agriculture	Many	Standardized	None	Low	None
Monopolistic competition	Retail trade	Many	Differentiated	Some	Low	Advertising and product differentiation
Oligopoly	Computers, oil, steel	Few	Standardized or differentiated	Some	High	Advertising and product differentiation
Monopoly	Public utilities	One	Unique product	Considerable	Very high	Advertising

differentiation gives managers the power to overrule the market price. Managers of oligopolies that produce differentiated products also tend to rely heavily on nonprice competition, whereas managers of oligopolies that produce nondifferentiated products do not. For example, computer firms try to increase their sales by building better computers and by advertising, whereas steel companies do little advertising. Monopolists also engage in advertising and public relations, although these activities are directed not at capturing the sales of other firms in the industry (no other firms exist) but rather at increasing total market demand and insulating the firm from the negative connotations sometimes associated with monopoly.

Table 7.1 summarizes many key features of each market structure. Be sure to look over this table before proceeding further. This chapter discusses perfect competition. Chapter 8 covers monopoly and monopolistic competition. Chapters 9 and 10 extend the monopoly model to consider sophisticated monopoly pricing strategies. Chapter 11 considers oligopoly.

MARKET PRICE IN PERFECT COMPETITION

In a perfectly competitive industry, market price, as we saw in Chapter 1, is determined by the intersection of the market demand and supply curves. The market demand curve shows the total amount that individual buyers of the commodity will purchase at any price; the market supply curve shows the total amount that individual suppliers of the commodity will supply at any price. Figure 7.1 shows

FIGURE 7.1**Determination of Price in a Perfectly Competitive Market**

Equilibrium price is \$10, and equilibrium quantity is 24,000 units.



the market demand and supply curves for a good produced in a perfectly competitive market. As is ordinarily the case, the market supply curve slopes upward to the right; that is, price increases generally result in higher industry output because managers find it profitable to expand production. Also, in accord with Chapters 1 and 2, the market demand curve slopes downward to the right; that is, price increases generally result in less product being demanded.

To determine the equilibrium price, which is the price that will eventually prevail in this market, we must find the price at which market supply equals market demand.¹ The demand curve in Figure 7.1 is

$$P = 22 - 0.5Q_D \quad (7.1)$$

where P is the price (in dollars) of this good and Q_D is the quantity demanded (in thousands of units). The supply curve in Figure 7.1 is

$$P = 4 + 0.25Q_S \quad (7.2)$$

where Q_S is the quantity supplied (in thousands of units). Because the equilibrium price is at the level where Q_D (the quantity demanded) equals Q_S (the quantity supplied),

$$\begin{aligned} P_D &= 22 - 0.5Q = 4 + 0.25Q = P_S \\ 0.75Q &= 18 \\ Q &= 24 \end{aligned}$$

1. Recall from Chapter 1 that an equilibrium price is a price that can be maintained. If conditions do not change, the actual price tends to equal the equilibrium price.

Substituting 24 for Q_D in equation (7.1), we find that $P = \$10$. (If we substitute 24 for Q_S in equation (7.2), we get the same result.) Therefore, as shown in Figure 7.1, price is expected to equal \$10, and output is expected to equal 24,000 units.

Although Figure 7.1 shows that both total quantity demanded and total quantity supplied depend on price, this does not mean an individual manager can affect price. According to the market demand curve in equation (7.1),

$$P = 22 - 0.5Q_D$$

If 1,000 firms are in this market, each produces, on the average, only 24 units of the product. Even if an individual firm doubles its output (from 24 to 48 units), the effect on price is minuscule. Specifically, an output increase of 24 units results in a price reduction of only 1.2 cents, or about 0.1%.² This means that managers in this market essentially face a horizontal demand curve. No matter how many units one manager sells, the market price remains the same. Whereas the demand curve for the output of the entire industry slopes downward to the right (as shown in Figure 7.1), the demand curve for the output of any single firm is regarded as horizontal (at a price of \$10 in this case).

SHIFTS IN SUPPLY AND DEMAND CURVES

Shifts in the market supply or demand curves result in price changes (recall Chapter 1). For example, if the supply curve in Figure 7.1 shifts to the left, the price is expected to rise. Shifts in market supply and demand curves have significant consequences for firm performance, and managers must try to anticipate them and respond as best they can.

For present purposes, managers must understand the factors affecting the supply and demand curves of the products they buy and sell. There is no need here to dwell at length on the factors causing shifts in demand curves; they have been discussed in Chapter 2. But it is worth recalling from Chapter 1 that two of the most important factors causing shifts in supply curves are technological advancements and changes in input prices. Improvements in technology tend to shift a product's supply curve to the right because they often permit managers to reduce their costs. On the other hand, increases in input prices tend to shift a product's supply curve to the left because they push up the firm's costs.

THE OUTPUT DECISION OF A PERFECTLY COMPETITIVE FIRM

How much output should managers at a perfectly competitive firm produce? As we discussed previously, managers in a perfectly competitive firm cannot affect the market price of their product, and they must sell any output (within their capabilities) at the market price. To illustrate the manager's situation, consider the example in Table 7.2. The market price is \$10 per unit, and the manager can

2. If output increases by 24 units, Q increases by 0.024 because Q is measured in thousands of units. If Q increases by 0.024, P falls by $0.5(0.024) = 0.012$, according to the demand curve in equation (7.1). P is measured in dollars, so this amounts to 1.2 cents.

TABLE 7.2

Cost and Revenues of a Perfectly Competitive Firm

Units of Output Period	Price (Dollars)	Total Revenue (Dollars)	Total Fixed Costs (Dollars)	Total Variable Costs (Dollars)	Total Cost (Dollars)	Total Profit (Dollars)
0	10	0	1	0	1	−1
1	10	10	1	3	4	6
2	10	20	1	8	9	11
3	10	30	1	15	16	14
4	10	40	1	24	25	15
5	10	50	1	35	36	14
6	10	60	1	48	49	11
7	10	70	1	63	64	6
8	10	80	1	80	81	−1
9	10	90	1	99	100	−10

produce as much as she chooses. Hence the firm's total revenue at various output rates is given in column 3 of Table 7.2. The firm's total fixed cost (1), total variable cost, ($2Q + Q^2$), and total cost ($1 + 2Q + Q^2$) are given in columns 4, 5, and 6 of Table 7.2. Finally, the last column shows the firm's total profit.

Figure 7.2 shows the relationship between total revenue and cost and output. The vertical distance between the total revenue and total cost curves is the profit at the corresponding output. Below one unit of output and above seven units of output, this distance is negative. Because the manager can sell either large or small volumes of output at the same price per unit, the total revenue curve is a straight line through the origin with a slope equal to the fixed price. (Specifically, $TR = P \cdot Q$; because the price is constant, total revenue is proportional to quantity.) Because a manager in a perfectly competitive firm takes the price as given, the slope of the total revenue is always the market price.

The firm's profit (π) is expressed as total revenue (TR) minus total cost (TC)

$$\pi = TR - TC$$

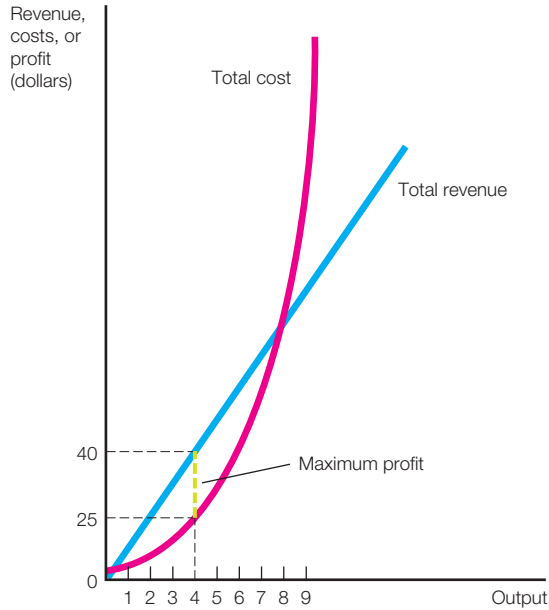
It follows that

$$\frac{\Delta \pi}{\Delta Q} = \frac{\Delta TR}{\Delta Q} - \frac{\Delta TC}{\Delta Q}$$

FIGURE 7.2

Relationship between Total Cost and Total Revenue of a Perfectly Competitive Firm

The output rate that would maximize the firm's profit is four units per time period. The profit (total revenue minus total cost) equals \$15.



If

$$\frac{\Delta \pi}{\Delta Q} = 0$$

then

$$\frac{\Delta TR}{\Delta Q} - \frac{\Delta TC}{\Delta Q} = 0$$

so

$$\frac{\Delta TR}{\Delta Q} = \frac{\Delta TC}{\Delta Q}$$

Here $\Delta TR/\Delta Q$ is the firm's marginal revenue. It represents a change in total revenue when the output changes by a small amount (usually $\Delta Q = 1$). The firm's total revenue is price times quantity or PQ . Therefore, marginal revenue is

QUANT OPTION

$$\pi = TR - TC$$

$$\frac{d\pi}{dQ} = \frac{dTR}{dQ} - \frac{dTC}{dQ}$$

To maximize profit,

$$0 = \frac{dTR}{dQ} - \frac{dTC}{dQ}$$

so

$$\frac{dTR}{dQ} = \frac{dTC}{dQ}$$

or

$$MR = MC$$

$\Delta TR/\Delta Q = P$. So the firm's marginal revenue is the product's price. This is not surprising. If the firm sells five units, its total revenue is $5P$; if it sells six units, its total revenue is $6P$; if it sells seven units, its total revenue is $7P$; and so on. Every time another unit is sold, total revenue rises by P . In this case $P = \$10$, so the firm's marginal revenue (always equal to price for a price taker) is $\$10$.

Consider the firm's total cost ($TC = 1 + 2Q + Q^2$). Therefore, $\Delta TC/\Delta Q = 2 + 2Q$, and $\Delta TC/\Delta Q$ is called the firm's marginal cost. It represents a change in the total cost (or variable cost) when output changes by a small amount (usually by $\Delta Q = 1$).

The condition $\Delta TR/\Delta Q = \Delta TC/\Delta Q$ is restated as $MR = MC$; that is, to maximize profit, the manager must set marginal revenue equal to marginal cost (if marginal cost is increasing). In the case of a price taker, the profit-maximizing condition reads $P = MC$ (because $P = MR$). This is why managers want to avoid these markets: The nature of competition is to grind the price down to marginal cost. The competitive pressure is relentless. There is no above-normal economic profit, nor should managers expect any (except in the short run). Clearly managers should never produce output where the marginal cost is greater than the marginal revenue. Hence

$$P = 10 = 2 + 2Q = MC \text{ or } Q = 4 \quad (7.3)$$

One other condition must exist for the $P = MC$ rule to yield a profit maximum. It must be the case that $\Delta(\Delta\pi/\Delta Q)/\Delta Q < 0$. Since $\Delta\pi/\Delta Q = P - MC$,

this means that $\Delta P/\Delta Q - \Delta MC/\Delta Q < 0$. Since $\Delta P/\Delta Q = 0$, then $-\Delta MC/\Delta Q < 0$ or $\Delta MC/\Delta Q > 0$. That is, the marginal cost must be increasing.

QUANT OPTION

For a maximum to occur, the sign of the second derivative of the profit function must be negative.

$$\begin{aligned}\frac{d^2\pi}{dQ^2} &= d\left(\frac{d\pi}{dQ}\right) \\ &= d\left(\frac{dTR}{dQ}\right) - d\left(\frac{dTC}{dQ}\right) \\ &= \frac{dMR}{dQ} - \frac{dMC}{dQ} < 0\end{aligned}$$

Because $MR = P$ and doesn't change as Q changes,

$$\frac{dMR}{dQ} = 0$$

Thus dMC/dQ must be positive because it has a negative sign in front of it and the whole equation must be negative.

Table 7.2, Figures 7.2 and 7.3, and equation (7.3) show that managers maximize the firm's profit at four units per time period. At this output level, the profit figure in the last column of Table 7.2 is the highest; the vertical distance between the total revenue and cost curves in Figure 7.2 is the largest; and the profit curve in Figure 7.3 is the highest.

QUANT OPTION

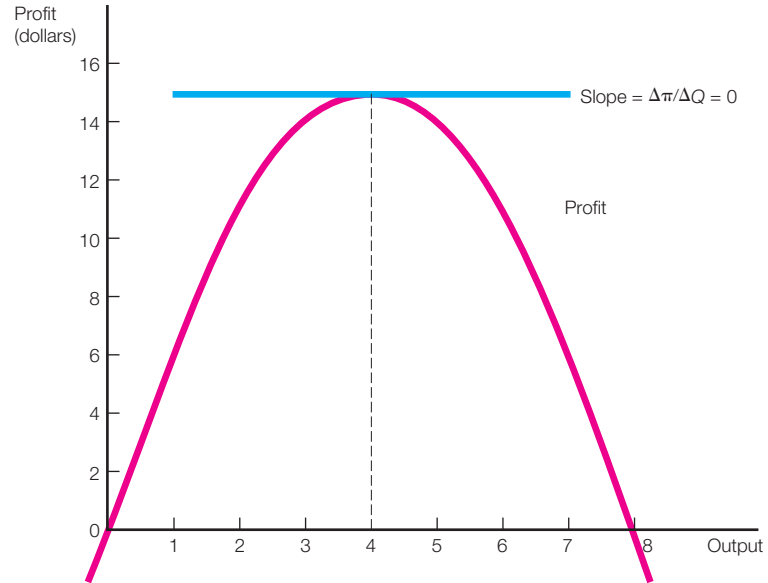
Of course the marginal revenue of the firm is $dTR/dQ = P$ because Q is not a function of P . The marginal cost is

$$dTC/dQ = 2 + 2Q$$

FIGURE 7.3

Relationship of Profit and Output of a Perfectly Competitive Firm

The output rate that maximizes profit is four units per time period. To maximize profit, the slope of the profit function ($\Delta\pi/\Delta Q$) must be zero.



It is worthwhile to present the marginal revenue and marginal cost curves as well as the total revenue and total cost curves. Table 7.3 shows the firm's marginal revenue and marginal costs at each output rate.

Figure 7.4 shows the resulting marginal revenue and marginal cost at each output rate. Because the manager takes the price as given, it is constant for all output levels ($P = MR$). Hence the marginal revenue curve is also the firm's demand curve, which (for the reasons discussed already) is horizontal.

The central point to note is that managers maximize profit at the output where the price (or marginal revenue) equals the marginal cost. Both the numbers in Table 7.3 and the curves in Figure 7.4 indicate that the price equals the marginal cost at an output rate of four units, which we know from Table 7.2, Figures 7.2 and 7.3, and equation (7.3) to be the profit-maximizing output.

SETTING THE MARGINAL COST EQUAL TO THE PRICE

Earlier we showed that if managers want to maximize firm value, they should set price equal to marginal cost when marginal cost is increasing. As we can see in Table 7.3 and Figure 7.4, this is true.³

3. Recall from Chapter 6 that in the short run, $MC = W/MP$, where W is the fixed wage of labor and MP is the marginal product of labor. Therefore, if the marginal cost is increasing, the marginal product of labor must be decreasing. A price-taking situation, therefore, implies diminishing marginal productivity of labor in the short run.

TABLE 7.3

Marginal Revenue and Marginal Cost: Perfectly Competitive Firm

Output per Period	Marginal Revenue	Marginal Cost ^a
0	10	2
1	10	4
2	10	6
3	10	8
4	10	10
5	10	12
6	10	14
7	10	16
8	10	18

^aThis column was calculated from $MC = 2 + 2Q$. This assumes that output can be sold and produced in noninteger, continuous amounts. Many goods meet this criteria—for example, gasoline, deli meats, and bulk agricultural products. Although these goods are priced on a per-unit basis (by the gallon, per pound), seldom do we purchase integer units. Other goods are produced and consumed only in integer units, such as cars, televisions, and compact disks. In this case, the marginal cost for output level n is calculated as the total cost at output level n minus the total cost for output level $n - 1$. For instance, the total cost at $Q = 3$ is 16, at $Q = 4$ is 25, at $Q = 5$ is 36 (see Table 7.2). Therefore, the marginal cost at $Q = 4$ is 9 and at $Q = 5$ is 11 (as opposed to 10 and 12 shown in this table). Under these conditions, the manager wants to produce the fourth unit because the marginal revenue from doing so (\$10) exceeds the marginal cost (\$9). However, the manager would not produce the fifth unit because the marginal cost of doing so (\$11) exceeds the marginal revenue (\$10). Therefore, the manager produces four units in the integer output case and in the continuous output case.

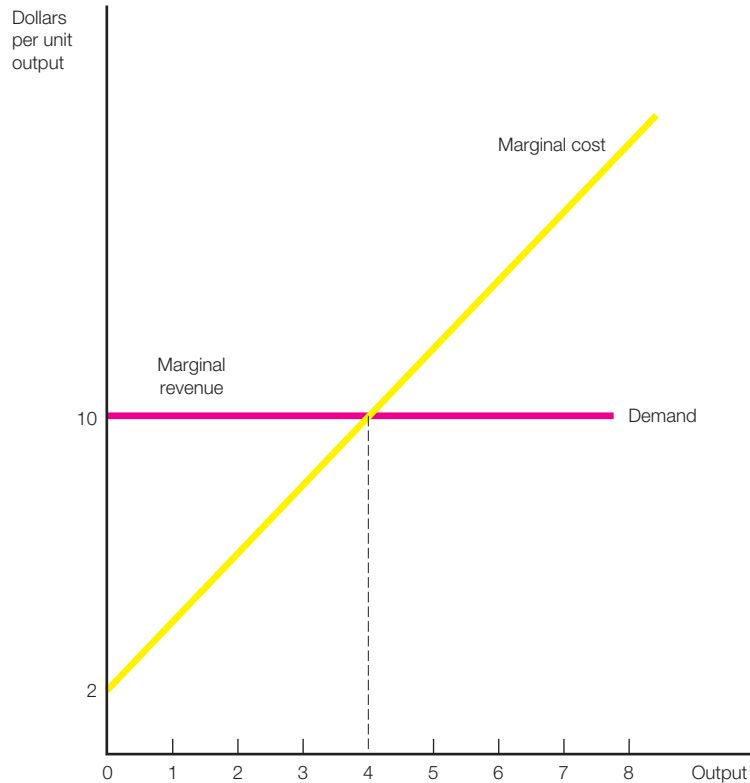
Managers in perfectly competitive markets often accrue negative profits, even if they satisfy the preceding rules ($P = MC$ and MC increasing). If the price is P_2 in Figure 7.5, the short-run average total cost exceeds the price at all possible outputs. Because the short run is too short (by definition) to permit the manager to alter the scale of plant, all she can do is produce at a loss or discontinue production. The decision to close a plant should answer one question: Does the product's price cover the average variable costs? For any output where price exceeds average variable costs, managers should produce, even though the price does not cover average total costs. If there is no output rate at which price exceeds the average variable cost, the manager is better off shutting the plant. Hence if the average variable cost curve is as shown in Figure 7.5, the manager will produce if the price is P_2 but not if it is P_1 .

It is essential to recognize that if managers shut a plant, they still incur fixed costs. Therefore, if the loss from producing is less than the firm's fixed costs (the loss of shutting down), it is more profitable (a smaller deficit) to produce than

FIGURE 7.4

Marginal Revenue and Marginal Cost of a Perfectly Competitive Firm

When output is at the profit maximizing level of four units, price (= marginal revenue) equals marginal cost.



to discontinue production. Another way of expressing this is if the loss per unit of output is less than the average fixed cost—that is, if $ATC - P < AFC$, where ATC is average total cost, P is price, and AFC is average fixed cost. This is true if $ATC < AFC + P$ because P is merely added to both sides of the inequality. Subtracting AFC from both sides results in $ATC - AFC < P$. $ATC - AFC$ is the average variable cost; thus it is better to produce than to shut down production if the price exceeds the average variable cost.

Managers must manage the cash flow they control, which is total revenue (because managers control the Q of P^*Q). They must also manage variable cost because it is a function of Q . But the fixed cost is not part of controllable short-run

In the eastern part of the United States, (Delaware, Illinois, Indiana, Kentucky, Maryland, Michigan, New Jersey, North Carolina, Ohio, Pennsylvania, Tennessee, Virginia, West Virginia, and the District of Columbia), power can be transmitted among power companies on the PJM Interconnection (a regional transmission organization). There are 650 member companies and 1,325 generating stations on the network. It is the largest competitive wholesale electric market in the world. Companies that have excess power can sell it to the highest bidder, and companies that require more power can buy power from the grid. The grid operator has an independent monitor, Monitoring Analytics. The monitor reports that PJM's prices were set by marginal units operating at or near their marginal cost. Joseph Bowring, president of Monitoring Analytics, reports that "this is evidence of competitive behavior and competitive market outcomes." That's what the theory tells us.

Historically, the international iron ore market tended to operate on an annual price that was reset every year. Customers and suppliers signed contracts based on this rate, despite the fact that demand or supply conditions arising during the course of that year could cause more to be demanded at that price than supplied or more supplied than demanded. Prices were not free to go up in the former case or down in the latter case. But over time, a spot market developed reflecting the market conditions in supply and demand in real time.

Marius Kloppers, head of BHP Billiton, is a major advocate of the spot market. Kloppers notes that spot pricing "will mean that the cost of the highest marginal tonne (read highest seller's reservation price) in any particular system, be it iron ore, coking or thermal coal, or manganese, will be more efficiently transferred into the global price. And that means, over the longer haul, the lowest-cost operators (those with the lowest seller's reservation price) will secure both a greater share of overall demand and better prices." With respect to exports to China, Kloppers notes that true market prices encourage

Australian producers to expand more vigorously (as demand is growing) and that the new Australian tonnes to meet the expanded Chinese demand replace the most expensive Chinese domestic production. "And each time someone (a high marginal cost producer) is squeezed out, the highest marginal cost is lowered and thus the reference price falls too." Ah, the beauty of price equals marginal cost in action.

Kloppers appears again in a bid to acquire Potash Corp of Saskatchewan, Canada. If successful, his strategy will be to operate the mines at full capacity. Because they are low-cost mines, Kloppers believes that such a move would lower prices in the short run. However, he believes that such production will lead to higher prices over the long run "as rival marginal high-cost production is deferred."

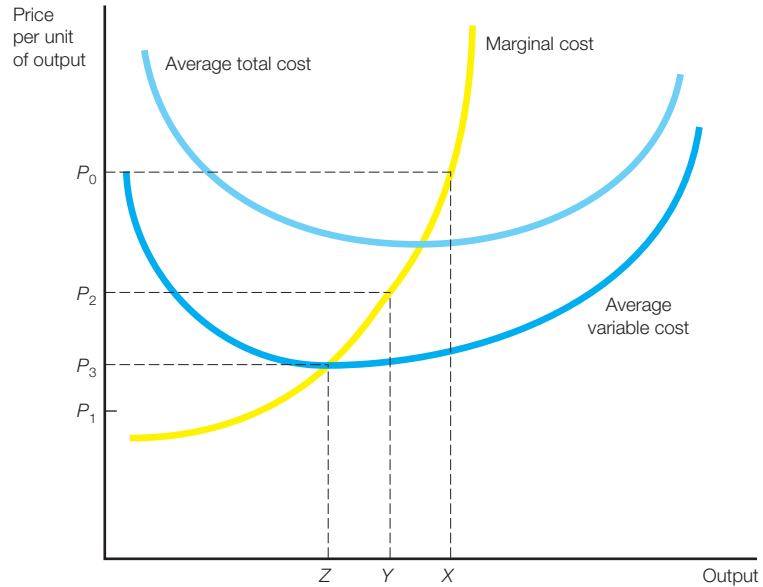
The market share of the big three cement producers in central Africa (initially 93%) has fallen by 14% as new producers have entered the market with cheaper prices. The new entrants' factories have lower maintenance costs and higher energy efficiency. In addition, they have lower clinker (a mineral element that goes into cement production) costs. Finally, the new entrants are already in the building supplies industry and can add the cement into their existing distribution network. These lower costs are passed on to consumers, and the lower marginal cost has shifted the supply curve of cement downward yielding lower cement prices.

Sources: "PJM Wholesale Power Price Rises in 2010, but Markets Still Competitive, Monitor Says," *Platts Inside FERC*, August 16, 2010, at www.lexisnexis.com/Inacui2api/; Matt Stevens, "Shift to Rational Pricing Will End Annual Conflict," February 13, 2010, at www.theaustralian.com.au/business/opinion/shift-to-rational-pricing-will-end-annual-conflict/story-e6frg9if-1225829866038; Andy Stevens and Brenda Bouw, "Chinese Bid for Potash Corp Would Cut Revenues, BHP Warns," *The Globe and Mail*, September 21, 2010, at www.theglobeandmail.com/globe-investor/potash/chinese-bid-for-potash-corp-would-cut-revenues-bhp-warns/article1716883; Moses Michira, "New Cement Firms Ride on Low Prices to Grow Sales," *Business Daily*, January 27, 2011, at www.businessdailyafrica.com/New-cement-firms-ride-on-low-prices-to-grow-sales/-/539552/1096296/-/oemr6i-index.html.

FIGURE 7.5

Short-Run Average and Marginal Cost Curves

If the price is P_0 , the firm will produce an output of X ; if price is P_2 , it will produce an output of Y ; and if price is less than P_3 (that is, when P_3 equals the minimum of average variable cost), the firm will produce nothing.



cash flow. Thus managers want TR to exceed VC by as much as possible—that is, maximizing the controllable cash flow. If VC exceeds TR , they shouldn't produce because controllable cash flow is negative. If TR exceeds VC , they should produce because controllable cash flow is positive. Dividing both sides in the preceding sentence by Q yields $TR/Q = P > VC/Q = AVC$. Thus in addition to $P = MC$, P must exceed AVC .

The manager will be indifferent at P_3 , the price that is equal to the minimum point of average costs (output Z). The point (Z, P_3) is called the **shutdown point** because at that point, the price equals the minimum average variable cost. Price is also equal to marginal cost because it intersects with average variable cost at this point. The future is clear: At this price, the manager loses money equal to fixed cost if he produces, or he loses that money if he shuts down. At any price below P_3 , the manager shuts the plant. Therefore, the marginal cost curve (above the minimum average variable cost) is the supply curve of the firm; that is, if the price is P_2 , the firm produces Y , and if the price is P_0 , the firm produces X . These points (Y, P_2) and (X, P_0) are also points on the firm's marginal cost curve.

Shutdown point When the price equals the minimum average variable cost.

STRATEGY SESSION: Forecasting the Price of Salmon

Managers at a large diversified food firm set out to forecast the price of fresh salmon three years ahead. Such a forecast is needed because these managers must decide whether they should enter the business of supplying salmon. The firm's analysts estimate the quantity of fresh salmon to be supplied three years ahead. Because of substantial plans to expand the production of farmed Atlantic and Pacific salmon in Canada, Chile, Japan, and Ireland, this projected supply is considerably greater than the actual current supply. In addition, the analysts estimate the quantity of fresh salmon that is demanded three years ahead. Their results show that if the price of salmon remains unchanged over the next three years, the quantity supplied will exceed the quantity demanded by about 15% at the end of the three-year period.

The firm's analysts also estimate the price elasticity of demand for fresh salmon to be about -1.5 . This estimate, too, is based on the techniques

described in previous chapters. Like the other estimates presented here, it is regarded as rough but useful.

Before they can determine whether to enter the salmon market, managers need to estimate the future price of salmon. (The firm's analysts believed that the quantity supplied three years hence will be approximately equal to their estimates, regardless of whatever changes occur in price during this three-year period.) Using the previous estimates, they envision that the quantity demanded will increase by 15% to reduce the gap between quantity supplied and quantity demanded. Because they estimate that the price elasticity of salmon is -1.5 , a 10% decrease in price will increase the quantity demanded by about 15%.

Note: This section is based on an actual case, although the numbers and situation have been disguised somewhat.

PROBLEM SOLVED: The Kadda Company

As an illustration, consider the Kadda Company, a perfectly competitive firm with the following total cost function

$$TC = 800 + 6Q + 2Q^2$$

where TC is the total cost (in dollars) and Q is the firm's output per day. The firm's marginal cost is therefore

$$\Delta TC / \Delta Q = MC = 6 + 4Q$$

If the price of Kadda's product equals \$30, the manager should set output so that

$$MC = \Delta TC / \Delta Q = 6 + 4Q = 30 = P \quad (7.4)$$

In other words, the manager should set marginal cost equal to price (\$30). Solving equation (7.4) for Q , we find the manager should set output equal to six units per day. To make sure the price is not less than average variable cost at that output, we note that because the firm's total variable cost equals $6Q + 2Q^2$, its average variable cost (AVC) equals

$$AVC = (6Q + 2Q^2) / Q = 6 + 2Q$$

Therefore, if $Q = 6$, average variable cost equals $6 + 2(6)$ or \$18, which is less than the price of \$30.

QUANT OPTION

Kadda's marginal cost is

$$dTC/dQ = MC = 6 + 4Q$$

To summarize, if the manager maximizes profit or minimizes loss, the output is set so the short-run marginal cost equals the price and the marginal cost is rising. But this proposition has an exception: If the market price is below the firm's average variable costs at every output level, the manager minimizes loss by discontinuing production.

ANOTHER WAY OF VIEWING THE PRICE EQUALS MARGINAL COST PROFIT-MAXIMIZING RULE

If a firm has one fixed input (say capital) and one variable input (say labor), how much of its variable input should it utilize? This is an important question for managers of firms large and small. To answer it, we must define the marginal revenue product of the variable input and the marginal expenditure of the variable input.

The **marginal revenue product (MRP)** is the amount an additional unit of the variable input adds to the firm's total revenue. The input adds to total revenue because it allows managers to produce more output. Letting MRP_L be the marginal revenue product of the labor input,

$$MRP_L = \Delta TR / \Delta L \quad (7.5)$$

where ΔTR is the change in total revenue resulting from a change of L , the amount of labor input used by the firm. It can easily be proven that MRP_L equals labor's marginal product times the firm's marginal revenue. To see this, note that marginal revenue (MR) equals $\Delta TR / \Delta Q$, where ΔQ is the change in the firm's output, and that

$$MRP_L = \Delta TR / \Delta L = (\Delta TR / \Delta Q)(\Delta Q / \Delta L)$$

Because $\Delta Q / \Delta L$ equals labor's marginal product (MP_L), it follows that

$$MRP_L = (MR)(MP_L) \quad (7.6)$$

which is what we set out to prove. Let's view the intuition. If managers use ΔL more labor, they produce ΔQ more units of the firm's product—that is, the marginal product of labor. If managers take these additional ΔQ units to market, they will generate ΔTR in revenue (marginal revenue). The marginal revenue per unit

Marginal revenue product (MRP)

The amount an additional unit of the variable input adds to the firm's total revenue.

times the number of units gives the additional revenue obtained by managers as the result of using an additional unit of labor.

The marginal expenditure (ME) is the amount an additional unit of labor adds to the firm's total costs. That is, letting ME_L be the marginal expenditure on labor,

$$ME_L = \Delta TC / \Delta L \quad (7.7)$$

where ΔTC is the change in total cost resulting from a change in the amount of labor, ΔL . If managers can buy all the labor they want at the price of \$10 per unit, ME_L equals \$10. In some cases, however, managers must pay a higher price for labor to get more of it; in such cases, ME_L exceeds the price of labor (as we will show in the monopsony section of Chapter 8). Note that $\Delta TC / \Delta L$ can be written as $(\Delta TC / \Delta Q)(\Delta Q / \Delta L)$, where $\Delta TC / \Delta Q$ is the change in the firm's total cost (ΔTC) divided by the firm's change in output (ΔQ). $\Delta TC / \Delta Q$ is the firm's marginal cost (MC), or the change in the firm's total cost as its output is changed by a small amount.

QUANT OPTION

In more technical terms,

$$MRP_L = dTR/dL$$

and

$$ME_L = dTC/dL$$

To maximize profit, managers should use labor where its marginal revenue product equals its marginal expenditure. In other words, managers should set

$$MRP_L = ME_L \quad (7.8)$$

Again, let's view the intuition. To maximize profit, managers need to expand any activity as long as the marginal benefit exceeds the marginal cost. They should stop expanding it when the marginal benefit (in this case MRP_L) equals the marginal cost (in this case ME_L). To generalize this further, rewrite equation (7.8) as

$$\left(\frac{\Delta TR}{\Delta Q} \right) \left(\frac{\Delta Q}{\Delta L} \right) = \left(\frac{\Delta TC}{\Delta Q} \right) \left(\frac{\Delta Q}{\Delta L} \right)$$

or

$$\frac{\Delta TR}{\Delta Q} = \frac{\Delta TC}{\Delta Q}$$

or

$$MR = MC$$

We have again verified one of the most important rules of managerial economics: Managers should stop expanding output when marginal revenue equals marginal cost.

In the case of perfect competition, $MR = P$. So our rule becomes $P \times MP_L = ME_L = P_L$, where P_L is the price of a unit of labor. Consider the intuition once again. Hiring another unit of labor costs managers P_L . That laborer generates MP_L additional output, which, when managers take it to market, generates $P \times MP_L$ additional revenue for the firm. Managers should continue to hire more labor as long as $P \times MP_L > P_L$, and they won't hire labor if $P \times MP_L < P_L$. The stopping rule to maximize profit is $P \times MP_L = P_L$. Dividing both sides by MP_L gives $P = P_L/MP_L$. As was shown in Chapter 6, $P_L/MP_L = MC$. Thus to maximize profit in the perfectly competitive market, managers should expand production until $P = MC$.

PRODUCER SURPLUS IN THE SHORT RUN

In Chapter 3 we examined consumer surplus and saw that it equals the difference between the market price and the price consumers are willing to pay (their reservation price). Now we introduce a parallel idea, called producer surplus, from the supply side of the market. **Producer surplus** is the difference between the market price and the price the producer is willing to receive for a good or service (the producer's reservation price). As we showed in our analysis of a perfectly competitive market, a producer's reservation price is the marginal cost of producing a good or service (above the break-even point of the firm). Panel A of Figure 7.6 shows this surplus in the shaded area. The firm's profit before accounting for fixed cost (its variable cost profit $P^*BC'D' = P^*BE$) is just total revenue (P^*Q^*) minus variable cost ($D'C'Q^*O$). But the variable cost is also just the area under the marginal cost up to output Q^* ; that is, EBQ^*O . This variable-cost profit is also the shaded area in Panel A of Figure 7.6; that is, P^*BE . Note the difference between profit (P^*BCD) and producer surplus. To arrive at a producer surplus, managers subtract only the variable cost from the total revenue, whereas to calculate their profit, they subtract both the fixed and variable costs from total revenue. Hence the variable-cost profit is larger than profit (by the level of fixed cost), and producer surplus and variable-cost profit are the same. Because the perfectly competitive firm's marginal cost represents its supply curve, we can view producer surplus as the difference between the supply curve and the price received for the good (area B in Panel B of Figure 7.6).

Figure 7.7 shows the market equilibrium. Just as market demand is the horizontal summation of individuals' demand curves for the product, market supply

Producer surplus The difference between the market price and the price the producer is willing to receive for a good or service (the producer's reservation price).

FIGURE 7.6

Producer Surplus and Variable-Cost Profit

Producer surplus for the firm is its variable-cost profit, or total revenue minus variable cost. Producer surplus for the market is the area above the supply curve but below the price received for the good because the supply curve is the horizontal summation of the competitive firms' marginal cost curves.

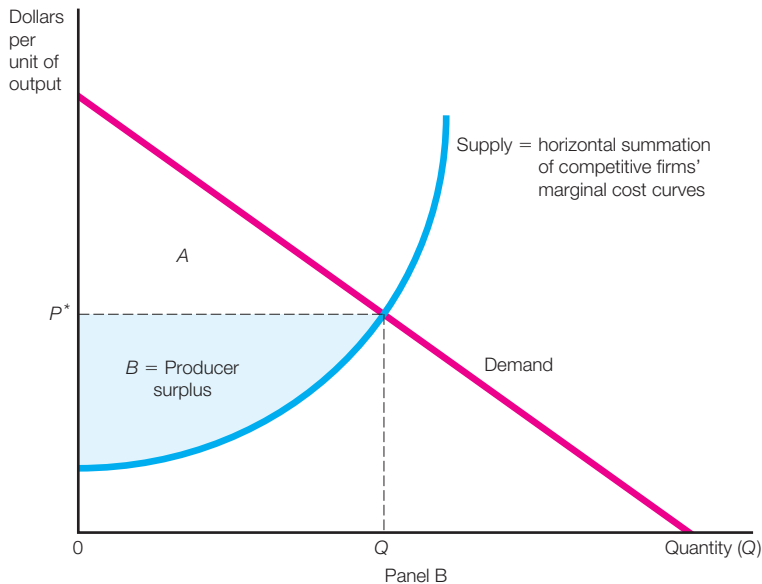
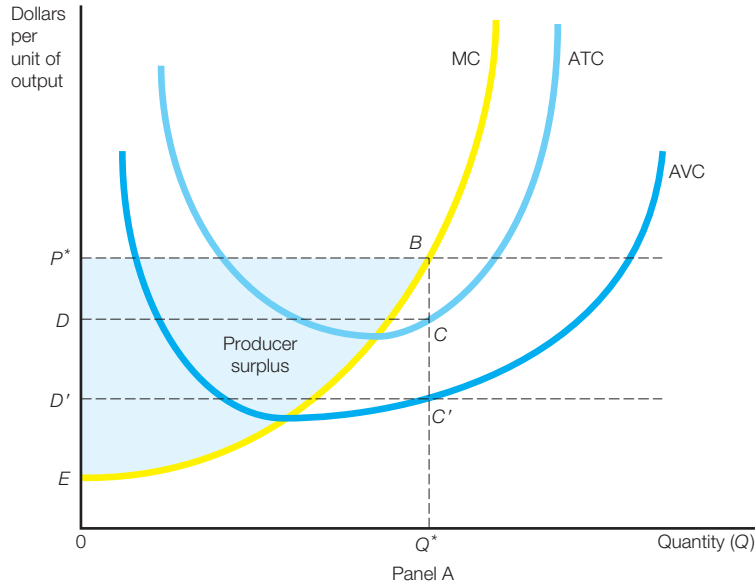
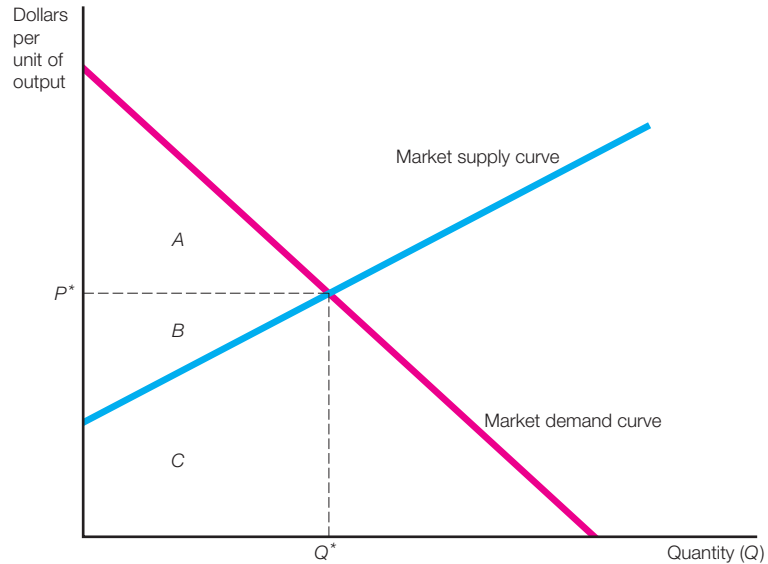


FIGURE 7.7

Market Social Welfare ($A + B$) of a Perfectly Competitive Price Policy, P^*

Social welfare at a given price (P^*) is measured by the sum of the consumer surplus (A) and the producer surplus (B).



is the horizontal summation of individual firms' supply curves for the product. Using the results for consumer surplus, we can see from Figure 7.7 that the market equilibrium price of P^* yields a consumer surplus of A and a producer surplus of B .

The sum of A and B , the total surplus, is the economist's measure of social welfare at the price P^* and the quantity Q^* . To understand this idea, think about the total benefit and cost that consumers and producers assign to the goods in the market. For consumers, this is the area beneath the demand curve, left of the equilibrium quantity—that is, the total amount consumers are willing to pay for the goods (areas A , B , and C in Figure 7.7). For producers, the total variable cost of supplying quantity Q^* is the area beneath the supply curve and left of the equilibrium quantity (area C in Figure 7.7). In the market consumers pay and producers receive P^* . Yet P^* is less than the total benefit and greater than the variable cost of the goods. In this sense the market exchange generates value for participants, represented by consumer, producer, and total surplus. In this case the difference between what the demanders are willing to spend (A , B , and C) and what the suppliers are willing to receive (C) is the measure of social welfare—that is, $A + B$.

Clearly the magnitude of the total surplus and its distribution between consumers and producers depend on the shape of the demand and supply curves. For instance, keeping the equilibrium at P^* , Q^* , a more gently sloped supply curve reduces producer surplus, whereas a more gently sloped demand curve reduces consumer surplus. But the thing to remember is that market exchanges, generally speaking, provide opportunities for gains; and as we see later, savvy managers can devise ways to capture a greater share of those gains for their firms.

We use this measure of social welfare to show the rationale of antitrust policy and the gains from trade in Chapter 17.

LONG-RUN EQUILIBRIUM OF THE FIRM

In the long run, how much will managers of a competitive firm produce? The long-run equilibrium position of the firm is at the point where its long-run average total cost curve⁴ equals the price. If the price exceeds the average total cost, economic profit is earned and new firms enter the industry. This increases supply, thereby driving down price and hence profit. If the price is less than the

4. This is also called the *long-run average cost curve*. Because all costs are variable in the long run, there is no need for an adjective in front of average costs, as there is in the short run, to distinguish between average total, average variable, and average fixed costs. There are only average costs in the long run.

PROBLEM SOLVED: Output at the Bergey Company

For example, suppose the Bergey Company's long-run average cost curve is

$$AC = 200 - 4Q + 0.05Q^2 \quad (7.9)$$

where AC is long-run average cost (in dollars) and Q is the firm's output per day. Because the Bergey Company operates in a perfectly competitive market, its output in the long run equals the value of Q that minimizes AC . Note from Figure 7.8 that the slope of the long-run average total cost curve when it is a minimum is $\Delta AC / \Delta Q = 0$; that is, GG' is tangent to AC at the output level V in Figure 7.8.

Forming $\Delta AC / \Delta Q$ from equation (7.9) and setting it = 0 gives $Q = 40$. Therefore, if managers at Bergey maximize profit, in the long run they will maintain an output of 40 units per day.

As indicated previously, the average cost equals the marginal cost at this output. To see this, note that because total cost equals Q times AC ,

$$\begin{aligned} TC &= Q(200 - 4Q + 0.05Q^2) \\ &= 200Q - 4Q^2 + 0.05Q^3 \end{aligned}$$

where TC is total cost.

The firm's marginal cost is $MC = \Delta TC / \Delta Q$. Therefore, $MC = \Delta TC / \Delta Q = 200 - 8Q + 0.15Q^2$. Because $Q = 40$,

$$MC = 200 - 8(40) + 0.15(40)^2 = 120$$

Also, inserting 40 for Q in equation (7.9),

$$AC = 200 - 4(40) + 0.05(40)^2 = 120$$

Therefore, marginal cost equals average cost when $Q = 40$. (Both marginal cost and average cost equal \$120. This means that the long-run equilibrium price is \$120.)

average total costs for any firm, that firm will exit the industry. As firms exit, supply falls, causing price and profit to rise. Only when economic profit is zero (which means that long-run average cost equals price) is a firm in long-run equilibrium.

Recall from Chapter 1 that economic profit is not the same as accounting profit. Economic profit is profit above and beyond what the owners could obtain elsewhere from the resources they invest in the firm. Therefore, long-run equilibrium occurs when the owners receive no more (and no less) than they could obtain elsewhere from these resources.

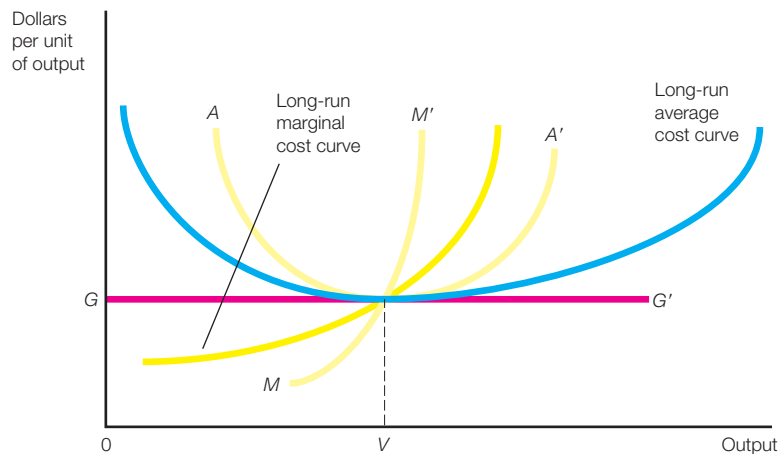
More specifically, the price must equal the *lowest value* of the long-run average total cost. That is, in equilibrium managers produce at the minimum point on their long-run average cost curves. To see why, note that if managers maximize their profit, they must operate where price equals long-run marginal cost. Also, we just saw that they must operate where price equals long-run average cost. If both of these conditions are satisfied, it follows that long-run marginal cost must equal long-run average cost. And we know from Chapter 6 that long-run marginal cost is equal to long-run average cost only at the point at which long-run average cost is a minimum. Consequently this point is the equilibrium position of the firm.

To illustrate this equilibrium position, consider Figure 7.8. When all the adjustments are made, price equals G . Because price is constant, the demand curve

FIGURE 7.8

Long-Run Equilibrium of a Perfectly Competitive Firm

In long-run equilibrium, the firm produces an output of V , and price = marginal cost (both long-run and short-run) = average cost (both long-run and short-run).



is horizontal, and therefore the marginal revenue curve is the same as the demand curve, both being GG' . The equilibrium output of the firm is V , and its optimally sized plant is described by the short-run average and marginal cost curves AA' and MM' . At this output and with this plant, we see that long-run marginal cost equals short-run marginal cost equals price. This ensures that the manager maximizes profit. Also, long-run average cost equals the short-run average cost equals price; this ensures that economic profit is zero. Because long-run marginal cost and long-run average cost must be equal, the equilibrium point is at the bottom of the long-run average cost curve.

THE LONG-RUN ADJUSTMENT PROCESS: A CONSTANT-COST INDUSTRY

Having looked at the behavior of managers at a perfectly competitive firm in the short and long runs, we turn to the long-run adjustment process of a perfectly competitive industry. We assume that this industry is a *constant-cost industry*, which means that expansion of the industry does not increase input prices. Figure 7.9 shows long-run equilibrium under conditions of constant cost. The top panel shows the short-run and long-run cost curves of a typical firm in the industry. The bottom panel shows the demand and supply curves in the market as a whole, D being the original demand curve and S the original short-run supply curve. We assume the industry is in long-run equilibrium, with the result that the price (\$6 per unit) equals the minimum value of the long-run (and short-run) average cost.

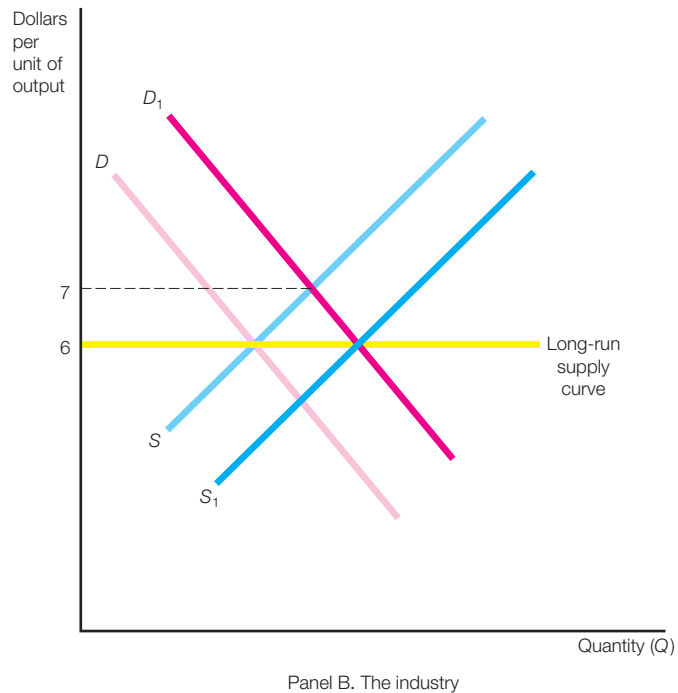
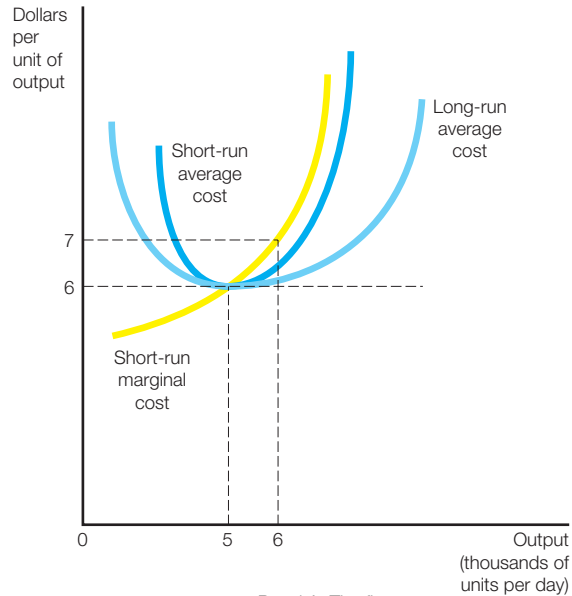
Suppose now that the demand curve shifts to D_1 . In the short run, with the number of firms fixed, the product price rises from \$6 to \$7 per unit; each firm expands its output from 5,000 to 6,000 units per day; and each firm makes economic profit because the new price, \$7, exceeds the short-run average costs of the firm when the output is 6,000 units per day. The result is that firms enter the industry and the supply curve shifts to the right. In a constant-cost industry, entrance of new firms does not influence the costs of existing firms. The inputs used by this industry are used by other industries as well, and new firms in this industry do not bid up the price of inputs and hence raise the costs of existing firms. Neither does the entry of new firms reduce existing firms' costs.

Hence a constant-cost industry has a horizontal long-run supply curve. Because output can be increased by increasing the number of firms producing 5,000 units of output per day at an average cost of \$6 per unit, the long-run supply curve is horizontal at \$6 per unit. So long as the industry remains in a state of constant costs, its output can be increased indefinitely. If price exceeds \$6 per unit, firms enter the industry; if price is less than \$6 per unit, firms leave the industry.

FIGURE 7.9

Long-Run Equilibrium in a Constant-Cost Industry

A constant-cost industry has a horizontal long-run supply curve, as shown in panel B. If demand shifts upward from D to D_1 , the consequent increase in price (to \$7 per unit) results in the entry of firms, which shifts the supply curve to the right (to S_1), thus pushing the price back to its original level (\$6 per unit).



Therefore long-run equilibrium can occur in this industry only when price is \$6 per unit. And industry output can be raised or lowered, in accord with demand conditions, without changing this long-run equilibrium price.

THE LONG-RUN ADJUSTMENT PROCESS: AN INCREASING-COST INDUSTRY

Not all industries are constant-cost industries. Next we consider the case of an *increasing-cost industry*, which occurs when industry expansion increases input prices.⁵ An increasing-cost industry is shown in Figure 7.10. The original conditions are the same as in Figure 7.9: In panel B, D is the original demand curve, S is the original supply curve, \$6 per unit is the equilibrium price, and LL' and AA' in the top panel are the long-run and short-run average cost curves of each firm. As in Figure 7.9, the original position is one of long-run equilibrium because price equals the minimum value of long-run (and short-run) average cost.

Assume now that the demand curve shifts to D_1 , with the result that the product price goes up and the firms earn economic profit, attracting new entrants. More and more inputs are needed by the industry, and in an increasing-cost industry, the prices of the inputs rise with the amount used by the industry. Therefore, the cost of inputs increases for established firms as well as entrants, and the average cost curves are pushed up to $L_1L'_1$ and $A_1A'_1$.

If each firm's marginal cost curve is shifted to the left by the increase in input prices, the industry supply curve tends to shift to the left. But this tendency is more than counterbalanced by the increase in the number of firms, which shifts the industry supply curve to the right. The latter effect *must* more than offset the former effect because otherwise there is no expansion in total industry output. (No new resources are attracted to the industry.) This process of adjustment must go on until a new point of long-run equilibrium is reached. In Figure 7.10 this point is where the price of the product is \$7 per unit and each firm produces 6,000 units per day.⁶

An increasing-cost industry has a positively sloped long-run supply curve. That is, after long-run equilibrium is achieved, increases in output require increases in the price of the product. For example, points X and Y in Figure 7.10 are on the long-run supply curve for the industry. The difference between constant-cost and increasing-cost industries is this: In constant-cost industries, new firms enter in response to an increase in demand until the price returns to its original level; whereas in increasing-cost industries, new firms enter until the minimum point on the long-run average cost curve has increased to the point where it equals the new, higher price.⁷

Finally, some industries are neither constant-cost nor increasing-cost industries: They are decreasing-cost industries. Their long-run supply curves

5. In addition to constant-cost and increasing-cost industries, there are also decreasing-cost industries, which are the most unusual case, although quite young industries may fall into the category. External economies, which are cost reductions that occur when an industry expands, may be responsible for the existence of decreasing-cost industries. For example, the expansion of an industry may improve transportation and reduce the costs of each firm in the industry. A decreasing-cost industry has a negatively sloped long-run supply curve.

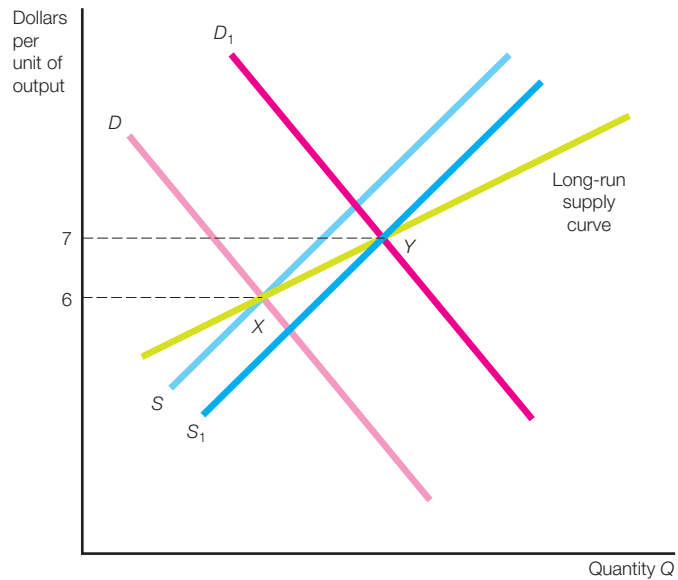
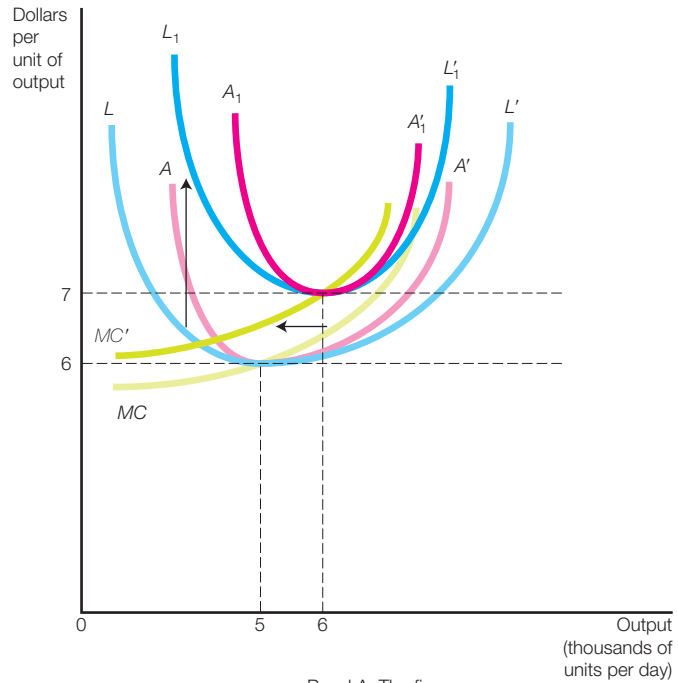
6. We cannot be sure that the firm's new output exceeds its old output as shown in Figure 7.10. It is possible for its new output to be less than or equal to its old output.

7. This is not the only way in which equilibrium can be achieved in increasing-cost industries. It is also possible that the increase in input prices (due to the expansion of industry output) raises average cost more than the increase in demand raises average revenue. Therefore, firms may experience losses, some may leave the industry, and the remaining firms may produce more.

FIGURE 7.10

Long-Run Equilibrium in an Increasing-Cost Industry

An increasing-cost industry has a positively sloped long-run supply curve, as shown in panel B. After long-run equilibrium is achieved, increases in output require increases in the price of the product.



are negatively sloped. For further discussion of these industries, which are encountered less frequently than constant-cost or increasing-cost industries, see footnote 5.

HOW A PERFECTLY COMPETITIVE ECONOMY ALLOCATES RESOURCES

It is important for managers to understand how a competitive economy allocates resources. Without such an understanding, they cannot interpret or anticipate the fundamental changes that may occur. To illustrate this allocation process, we take a simple case: Consumers become more favorably disposed toward corn and less favorably disposed toward rice than in the past. What happens in the short run? The rising demand for corn increases its price and results in some increase in the output of corn. However, corn output cannot be increased substantially because the capacity of the industry cannot be expanded in the short run. Similarly, the falling demand for rice reduces its price and results in some reduction in the output of rice. But the output of rice cannot be curtailed greatly because firms continue to produce as long as they can cover their variable costs. Because of the increased price of corn and the decreased price of rice, corn producers earn economic profit and rice producers show economic loss. Producers reallocate resources to correct this imbalance.

When short-run equilibrium is achieved in both the corn and rice industries, the reallocation of resources is not yet complete because there has not been enough time for producers to build new capacity or liquidate old capacity. In particular, neither industry operates at minimum average cost. The corn producers may operate at greater than the output level where average cost is a minimum; and the rice producers may operate at less than the output level where average cost is a minimum.

What occurs in the long run? The shift in consumer demand from rice to corn results in greater adjustments in output and smaller adjustments in price than in the short run. In the long run, existing firms can leave rice production and new firms can enter corn production. As firms leave rice production, the supply curve shifts to the left, causing the price to rise above its short-run level. The transfer of resources out of rice production ceases when the price has increased and costs have decreased to the point where loss no longer occurs.

Whereas rice production loses resources, corn production gains them. The short-run profit in corn production stimulates the entry of new firms. The increased demand for inputs raises input prices and cost curves in corn production, and the price of corn is depressed by the movement to the right of the supply curve because of the entry of new firms. Entry stops when economic profit is no longer being earned. At that point, when long-run equilibrium is achieved, more firms and more resources are used in the corn industry than in the short run.

SUMMARY

1. Managers of perfectly competitive firms set output levels so that price equals marginal cost. If there is an output level where price exceeds average variable cost, it pays for managers to produce in the short run, even though price does not cover average total costs. But if there is no output level where price exceeds average variable cost, managers are better off to produce nothing at all. In the long run, managers produce at the minimum point on their long-run average total cost curve. The price tends to be at the level where the market demand curve intersects the market supply curve. The short-run supply curve of a perfectly competitive firm is its marginal cost curve above the point of the minimum average variable cost.

2. Producer surplus is equivalent to the firm's variable-cost profit—that is, total revenue less variable costs. The producer surplus is the difference between the price a seller receives for its product and the seller's reservation price (the minimum price at which she would sell her product). This is a measure of welfare from a producer's perspective. When combined with the consumer surplus introduced in Chapter 3, the sum of the producer and consumer surpluses gives a measure of social welfare. We can use this measure to compare the benefits of different pricing proposals and the benefits of trade (as shown in Chapter 17).

3. A constant-cost industry has a horizontal long-run supply curve; an increasing-cost industry has a positively sloped long-run supply curve. If a constant-cost industry expands, there is no increase (or decrease) in input prices; if an increasing-cost industry expands, there is an increase in input prices.

wwwnorton.com/studyspace 

PROBLEMS

1. The Hamilton Company is a member of a perfectly competitive industry. Like all members of the industry, its total cost function is

$$TC = 25,000 + 150Q + 3Q^2$$

where TC is the firm's monthly total cost (in dollars) and Q is the firm's monthly output.

- a. If the industry is in long-run equilibrium, what is the price of the Hamilton Company's product?
 - b. What is the firm's monthly output?
2. In 2012, the box industry was perfectly competitive. The lowest point on the long-run average cost curve of each of the identical box producers was \$4, and this minimum point occurred at an output of 1,000 boxes per month. The market demand curve for boxes was

$$Q_D = 140,000 - 10,000P$$

where P was the price of a box (in dollars per box) and Q_D was the quantity of boxes demanded per month. The market supply curve for boxes was

$$Q_S = 80,000 + 5,000P$$

where Q_S was the quantity of boxes supplied per month.

- a. What was the equilibrium price of a box? Is this the long-run equilibrium price?
 - b. How many firms are in this industry when it is in long-run equilibrium?
3. The Burr Corporation's total cost function (where TC is the total cost in dollars and Q is quantity) is

$$TC = 200 + 4Q + 2Q^2$$

- a. If the firm is perfectly competitive and the price of its product is \$24, what is its optimal output rate?
 - b. At this output rate, what is its profit?
4. The supply and demand curves for pears are

$$Q_S = 10,000P$$

$$Q_D = 25,000 - 15,000P$$

where Q_S is the quantity (tons) supplied, Q_D is the quantity (tons) demanded, and P is the price per pear (in hundreds of dollars per ton).

- a. Plot the supply and demand curves.
 - b. What is the equilibrium price?
 - c. What is the equilibrium quantity?
5. The White Company is a member of the lamp industry, which is perfectly competitive. The price of a lamp is \$50. The firm's total cost function is

$$TC = 1,000 + 20Q + 5Q^2$$

where TC is total cost (in dollars) and Q is hourly output.

- a. What output maximizes profit?
 - b. What is the firm's economic profit at this output?
 - c. What is the firm's average cost at this output?
 - d. If other firms in the lamp industry have the same cost function as this firm, is the industry in equilibrium? Why or why not?
6. The long-run supply curve for a particular type of kitchen knife is a horizontal line at a price of \$3 per knife. The demand curve for such a kitchen knife is

$$Q_D = 50 - 2P$$

where Q_D is the quantity of knives demanded (in millions per year) and P is the price per knife (in dollars).

- What is the equilibrium output of such knives?
- If a tax of \$1 is imposed on each knife, what is the equilibrium output of such knives? (Assume the tax is collected by the government from the suppliers of knives.)
- After the tax is imposed, you buy such a knife for \$3.75. Is this the long-run equilibrium price?

EXCEL EXERCISE: PERFECT COMPETITION

Suppose that you are a price taker and the market price is given to you as 8 (from our market demand and supply exercise in Chapter 1). Your total cost is given as the total cost function determined in the Excel exercise on production and cost (Chapter 6), i.e., $TC = 5 + 2Q + 0.5Q^2$. From the text, we know that the profit-maximizing rule for a price taker is $P = MC$. From the above total cost formula,

$$MC = dTC/dQ = 2 + Q$$

(which assumes that the product can only be produced in integers)

the firm thus sets $P = 8 = 2 + Q = MC$

or $Q = 6$

The firm's total revenue is $TR = P \cdot Q = 8 \cdot 6 = 48$. The firm's total cost is $TC = 5 + 2Q + 0.5Q^2 = 5 + 2 \cdot 6 + 0.5 \cdot 6 \cdot 6 = 5 + 12 + 18 = 35$. The firm's profit is $\pi = TR - TC = 48 - 35 = 13$.

But suppose that you didn't know the above rule but instead knew that the market price was 8 and that the total cost relationship for your firm was

Quantity	Fixed Cost	Variable Cost	Total Cost
0	5	0.0	5.0
1	5	2.5	7.5
2	5	6.0	11.0
3	5	10.5	15.5
4	5	16.0	21.0
5	5	22.5	27.5
6	5	30.0	35.0
7	5	38.5	43.5
8	5	48.0	53.0
9	5	58.5	63.5

(continued)

Quantity	Fixed Cost	Variable Cost	Total Cost
10	5	70.0	75.0
11	5	82.5	87.5
12	5	96.0	101.0

Could you still determine the firm's profit-maximizing output and profit? Of course. Here's how with a spreadsheet:

Enter 0 in cell A1, 1 in cell A2, and so on, until you enter 12 in cell A13, replicating the quantity column in the table above.

Enter 5 in cell B1 and click on the lower-right-hand corner and drag down to cell B13. That should yield you a column of 5s, the firm's fixed cost.

Enter 0 in cell C1, 2.5 in cell C2, and continue until you enter 96 in cell C13. This gives you the firm's variable cost from the table.

Enter $=B1+C1$ in cell D1. Then click on the lower-right-hand corner and drag down to cell D13. This gives you the firm's total cost from the table.

You have now replicated the above table in the spreadsheet.

Enter $=D2-D1$ in cell E2, $=D3-D2$ in cell E3, and so on via the click-and-drag method. This will yield the firm's discrete marginal cost, which assumes that the firm can only produce output in integers.

For reference purposes, we will put the firm's continuous marginal cost (which assumes that the firm can produce output along a continuous spectrum—think gasoline production and gasoline sales) in column F. From above, the continuous marginal cost is $MC = 2 + Q$.

Enter $=2+A1$ in cell F1, $=2+A2$ in cell F2, and so on via click-and-drag. Notice that the continuous and discrete marginal costs are different, which reflects the different incremental changes in output. The continuous marginal cost exceeds the discrete marginal cost.

Enter 8 in cell G1. Click and drag this to cell G13. This will give you a column of 8s, the market price of the good.

Enter $=A1*G1$ in cell H1, $=A2*G2$ in cell H2, and so on, via click and drag. This will yield the firm's total revenue ($TR = P*Q$).

Enter $=H2-H1$ in cell I2, $=H3-H2$ in cell I3, and so on, via click and drag. This yields the firm's marginal revenue, i.e., its change in total revenue from making an additional sale.

Enter $=H1-D1$ in cell J1, $=H2-D2$ in cell J2, and so on, via click and drag. This yields a column of the firm's profits, $\pi = TR - TC$.

Now just search column J for the highest number. You should find 13 in the seventh row with the associated output of 6. Note that the continuous marginal cost (8) equals the price (8) at the output of 6 (as we showed at the beginning). You can also find the highest number by adding $=\text{Max}(J1:J13)$ in cell J14.

But we get the same results without the calculus. View the discrete output case. At output 6, the marginal revenue (price) is 8 and the discrete marginal cost is 7.5. Of course, the firm would want to produce the sixth unit, since cash flow in (8) exceeds cash flow out (7.5). But view the marginal revenue of the seventh unit (8) and the discrete marginal cost of producing the seventh unit (8.5). Clearly managers would not want to produce the seventh unit, because the cash flow in (8) is exceeded by the cash flow out (8.5).

Thus, both methods, the continuous (knowing calculus) and the discrete (with no calculus) yield the same solution.

MONOPOLY AND MONOPOLISTIC COMPETITION

The question faced by most managers is how to set prices and output when they have market power. As we will see, when managers possess market power, they have the ability to overrule the invisible hand described by Adam Smith. In these markets the equilibrium price set by the intersection of the supply and demand curves is rarely seen. We will first investigate this important issue by looking at how managers act when they have monopoly power. Managers with monopoly power do not have to consider the actions of market rivals because there are none. For example, US Airways is the only carrier flying between Ithaca, New York, and Philadelphia, Pennsylvania. In the winter Kubel's Restaurant is the only restaurant open in Barnegat Light, New Jersey. Only one supermarket may be open all night long in your area. The Philadelphia Gas Works is the only supplier of natural gas in Philadelphia, Pennsylvania.

The market demand curve for air travel between Ithaca and Philadelphia is the demand curve for US Airways. The market demand for winter restaurant meals in Barnegat Light is the demand curve seen by managers at Kubel's. Likewise, the market demand for overnight supermarket shopping and natural gas in Philadelphia are the demand curves facing those firm managers. Monopolies have no intramarket competition, and firm demand is equal to market demand.

LEARNING OBJECTIVES

Pricing and Output Decisions in Monopoly

Cost-Plus Pricing

Cost-Plus Pricing at Therma-Stent

Cost-Plus Pricing at Internet Companies and Government-Regulated Industries

Can Cost-Plus Pricing Maximize Profit?

The Multiple-Product Firm: Demand Interrelationships

Pricing of Joint Products: Fixed Proportions

Output of Joint Products: Variable Proportions

Monopsony

Monopolistic Competition

Advertising Expenditures: A Simple Rule

Using Graphs to Help Determine Advertising Expenditure

Advertising, Price Elasticity, and Brand Equity: Evidence on Managerial Behavior

Summary

Problems

Excel Exercise: Simple Monopoly

Appendix: Allocation of Output Among Plants

The demand faced by managers of monopolies is downward-sloping; that is, as price increases, quantity demanded decreases. Managers with market power face a pleasantly more complex decision relative to those in perfectly competitive markets. They must decide both price and quantity; they are no longer passive price takers. Relative to managers of perfectly competitive firms, they have more strategic power and are rewarded with higher economic profit. We now want to look at how managers maximize profit in such an environment.

Although being a monopolist comes with some degree of market power, it does not give managers *carte blanche*; they need to manage the demand characteristics of their product. If no one wants to fly between Ithaca and Philadelphia or shop at 3 A.M. or eat in a restaurant in the winter in Barnegat Light, the managers' monopoly power is virtually worthless. And even if managers create demand for their products, they still must efficiently manage costs and resources. Finally, monopolists still must worry about potential competitors. It is only 183 miles between Ithaca and Philadelphia, so driving or taking a bus is an option for many consumers. And though Barnegat Light has only one winter restaurant, a restaurant exists in a town three miles away, and home-cooked meals are a substitute. Many grocery stores are open all day and in the early evening, so customers can easily shop at times other than 3 A.M. Consumers need not heat their homes or hot water with gas; many choose to use oil or electricity. Cross elasticities (see Chapter 2) can tell us what goods, locations, and times are substitutes for a "monopoly" product.

So even when there is no intramarket competition, managers must work hard if substitute products, locations, and times exist. Managers need to understand product, spatial, and temporal competition, or they can make serious mistakes. Also, the higher the profit, the more others will test your market defenses and try to enter your market. Finally, if managers do too good a job and generate what is viewed as excess profit, authorities may try to regulate their actions in some way.

In this chapter, we examine issues the monopolist manager must consider in choosing the optimal price and quantity combination. We show how market power changes the thought process of managers, though the decision rule remains to produce where marginal revenue equals marginal cost. This simple rule dictates managerial behavior in all market structures.

We also show the profit-maximizing rule for managers in monopolistic competitive markets. In these markets managers still have market power, but they must deal with intramarket rivals. Although managers still face a downward-sloping demand curve, a lack of entry barriers allows others into the market. The world is more complicated; managers must consider the actions of these rivals in choosing their optimal strategy. Industries such as shirt manufacturing approximate monopolistic competition.

PRICING AND OUTPUT DECISIONS IN MONOPOLY

The monopolist behaves differently than the perfect competitor of Chapter 7. An unregulated monopolist maximizes profit by choosing the price and output where the difference between total revenue and total cost is the largest. For example, consider a monopolist with a demand curve of

$$P = 10 - Q$$

where P is the price per unit of the product and Q is the number of units demanded at that price. The monopolist has a total cost curve of

$$TC = 1 + Q + 0.5Q^2$$

The monopolist's total revenue is $TR = PQ$, or

$$TR = (10 - Q)Q = 10Q - Q^2$$

The total revenue and total cost for the monopolist at various levels of output are shown in Table 8.1. The manager maximizes profit at the output where total revenue exceeds total cost by the greatest amount. Figures 8.1 and 8.2 show the situation graphically.

TABLE 8.1

Cost, Revenue, and Profit of a Monopolist

Output	Price (Dollars)	Total Revenue (Dollars)	Variable Cost (Dollars)	Total Cost (Dollars)	Total Profit (Dollars)	Variable- Cost Profit (Dollars)
0	10	0	0	1	-1	0
1	9	9	1.5	2.5	6.5	7.5
2	8	16	4	5	11	12
3	7	21	7.5	8.5	12.5	13.5
4	6	24	12	13	11	12
4.5	5.5	24.75	14.625	15.625	9.125	10.125
5	5	25	17.5	18.5	6.5	7.5
6	4	24	24	25	-1	0
7	3	21	31.5	32.5	-11.5	-10.5
8	2	16	40	41	-25	-24
9	1	9	49.5	50.5	-41.5	-40.5
10	0	0	60	61	-61	-60

FIGURE 8.1

Total Revenue, Total Cost, and Total Profit of a Monopolist

To maximize profit, the monopolist chooses an output rate of three units per period of time and a price of \$7.

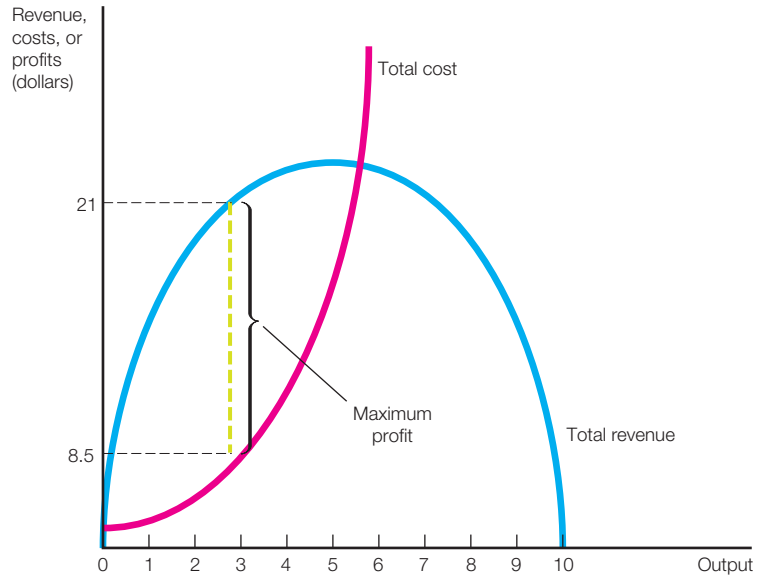
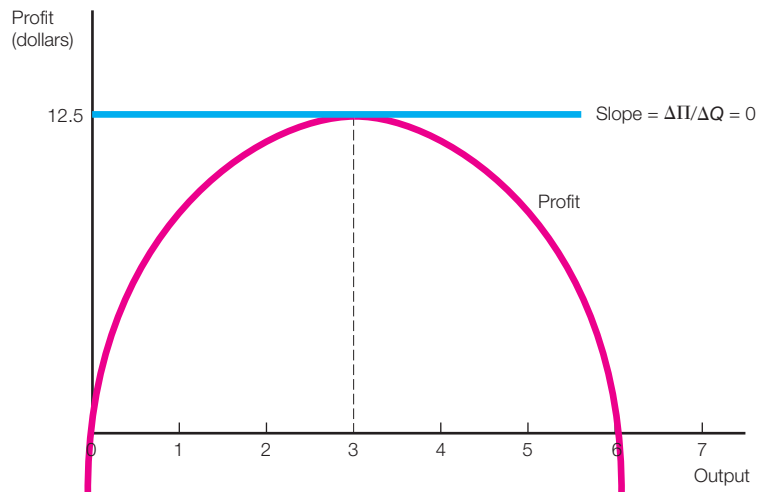


FIGURE 8.2

Profit and Output of a Monopolist

To maximize profit, the monopolist chooses an output rate of three units per time period and makes a profit of \$12.5.



Under monopoly, as under perfect competition, managers maximize profit if they set output at the point where marginal cost equals marginal revenue. As can be seen in Figure 8.2, profit is maximized when $\Delta\pi/\Delta Q = 0$. Recall from Chapter 7 that $\pi = TR - TC$ (that is, profit equals total revenue minus total cost) and that

$$\frac{\Delta\pi}{\Delta Q} = \frac{\Delta TR}{\Delta Q} - \frac{\Delta TC}{\Delta Q} = 0$$

which implies that $MR - MC = 0$ or that $MR = MC$.

Let us investigate the situation in Table 8.1 and Figures 8.1 and 8.2 in greater detail. The marginal revenue $= MR = \Delta TR/\Delta Q = 10 - 2Q$.

With a linear demand curve, the marginal revenue curve has the same dollar intercept as the demand curve (10) but it falls at twice the speed; that is, the marginal revenue curve has twice the slope of the demand curve. This is always true for linear demand curves (which we almost always use for illustrative purposes—of course, in the real world, demand curves could take any form that shows an inverse relationship of price and quantity). Therefore, a demand curve of $P = 250 - 12.5Q$ has a corresponding marginal revenue curve of $MR = 250 - 25Q$.¹

The total cost function is $TC = 1 + Q + 0.5Q^2$. Therefore, marginal cost $= MC = \Delta TC/\Delta Q = 1 + Q$.

Setting marginal revenue equal to marginal cost gives

$$MR = 10 - 2Q = 1 + Q = MC, \text{ or } Q = 3 \quad (8.1)$$

and hence $P = 10 - 3 = \$7$.

Unlike firms in a perfectly competitive market, the firm's marginal revenue is no longer constant; nor is it equal to price. Recall from Chapter 2 that

$$\begin{aligned} MR &= P \left[1 + \left(\frac{1}{\eta} \right) \right] \\ &= P \left[1 - \left(\frac{1}{|\eta|} \right) \right] \\ &= P - \left(\frac{P}{|\eta|} \right) \end{aligned} \quad (8.2)$$

QUANT OPTION

The monopolist's marginal revenue is

$$dTR/dQ = MR = 10 - 2Q$$

The monopolist's marginal cost is

$$dTC/dQ = MC = 1 + Q$$

1. Total revenue would be $TR = (250 - 12.5Q)Q = 250Q - 12.5Q^2$ and so $MR = dTR/dQ = 250 - 25Q$.

where MR is marginal revenue, P is price, and η is the price elasticity of demand. Because $\eta < 0$, the marginal revenue equals price minus $P/|\eta|$. Hence marginal revenue is price minus something positive—so price must exceed marginal revenue. In addition, no rational manager produces where marginal revenue is negative. (This implies that selling another unit decreases total revenue; because producing another unit would increase total costs, the manager could not maximize profit.) If managers are to produce where marginal revenue equals marginal cost, a negative marginal revenue implies a negative marginal cost. Total costs increase (not decrease) when managers increase production. If marginal revenue is positive, then from equation (8.2), $\eta < -1$ (that is, $|\eta| > 1$), which implies an elastic demand.² Thus a monopolist can not produce in the inelastic range of her demand curve if she is maximizing profit.

Table 8.2 and Figure 8.3 present the marginal revenue and marginal cost numbers for these functions; they substantiate that profit is maximized when marginal revenue equals marginal cost.

Note at the optimal output of three units (price = \$7), the demand is elastic (-2.33) and the marginal revenue is positive (\$4). It is also true that in a monopoly, price must exceed average variable cost if managers are to maximize profit (and from Table 8.1, we calculate at a production of three units, AVC is equal to \$2.5; that is, $AVC = VC/Q = \$7.5/3 = \2.5). If not, the monopolist is not covering variable cost and should shut the operation to reduce losses to only fixed cost.

2. From equation (8.2), if $|\eta| > 1$, then $1/|\eta| < 1$ and $[1 - (1/|\eta|)] > 0$. This makes $MR > 0$, because P must be positive.

STRATEGY SESSION: Why Monopoly Power Attracts Warren Buffett

Warren Buffett states that when he judges businesses, he does so based on their ability to raise prices way above the ability of the management team. Pricing power is his single most important criteria in evaluating a business. He says, "If you owned the only newspaper in town, up until the last five years or so, you had pricing power and you didn't have to go to the office." That's what a monopoly does: It gives the manager the ability to raise prices above marginal cost. He also says, "If you've got the power to raise prices without losing business to a competitor, you've got a very good business. And if you have to have a prayer session before raising the price by 10%, then you have a terrible business."

So monopoly is good. Buffett, of course, must also be sure that price covers average variable cost as we have shown in the text. Monopoly power without cost control is not a guaranteed ticket to success. Looking at the success of Berkshire Hathaway would suggest that Buffett is very good at covering all his costs.

Source: Andrew Frye and Dakin Campbell, "Buffett Says Pricing Power More Important Than Good Management," *Bloomberg News*, February 18, 2011, at www.bloomberg.com/news/2011-02-18/buffett-says-pricing-power-more-important-than-good-management.html.

TABLE 8.2

Marginal Cost and Marginal Revenue of a Monopolist

Price	Output	Marginal Cost ^a	Marginal Revenue ^b	Total Profit ^c	Elasticity
10	0	1	10	-1	$-\infty$
9	1	2	8	6.5	-9
8	2	3	6	11	-4
7	3	4	4	12.5	-2.33
6	4	5	2	11	-1.5
5.5	4.5	5.5	1	9.125	-1.22
5	5	6	0	6.5	-1
4	6	7	-2	-1	-0.67
3	7	8	-4	-11.5	-0.43
2	8	9	-6	-25	-0.25
1	9	10	-8	-41.5	-0.11
0	10	11	-10	-61	0

^aThe marginal cost is calculated from the equation $MC = 1 + Q$. This assumes that the product is produced in continuous amounts, like gasoline. If the product can be produced only in discrete amounts, like cars, the marginal cost for output n is defined as the total cost of producing n units minus the total cost of producing $n - 1$ units. Using the total cost information from Table 8.1, the marginal cost of producing two units is \$2.5 (that is, \$5 - \$2.5); the marginal cost of producing three units is \$3.5 (that is, \$8.5 - \$5); and the marginal cost of producing four units is \$4.5 (that is, \$13 - \$8.5). Why does this differ from the marginal cost of \$5 shown for output 4 in the table? Because the costs differ if you can produce continuously as opposed to only in discrete integer units.

^bThe marginal revenue is calculated from the equation $MR = 10 - 2Q$. This assumes that the product can be sold in continuous amounts, like gasoline. If the product can be sold only in discrete amounts, like cars, the marginal revenue

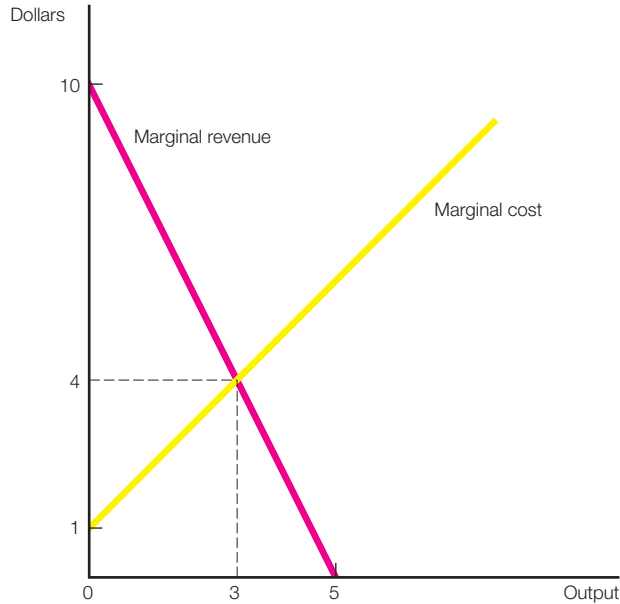
for output n is defined as the total revenue from selling n units minus the total revenue of selling $n - 1$ units. Using the total revenue information from Table 8.1, the marginal revenue of selling two units is \$7 (that is, \$16 - \$9); the marginal revenue of selling three units is \$5 (that is, \$21 - \$16); and the marginal revenue of selling four units is \$3 (that is, \$24 - \$21).

^cNote that using the discrete marginal revenue and the marginal cost gives the same result as the continuous analysis; that is, the profit-maximizing output is three units. In the discrete case, the firm would clearly produce the second unit because the marginal revenue exceeds the marginal cost (\$7 > \$2.5) and hence would increase profit. Likewise, it would produce the third unit because the marginal revenue exceeds the marginal cost (\$5 > \$3.5) and hence would increase profit. However, the firm would not produce the fourth unit because the marginal revenue is exceeded by the marginal cost (\$3 < \$4.5) and hence would decrease profit.

FIGURE 8.3

Marginal Revenue and Marginal Cost of a Monopolist

At the monopolist's profit-maximizing output (three units), the marginal cost equals the marginal revenue (at \$4).



It is easy to graphically show the price and output decision facing managers. Figure 8.4 shows the demand curve, the marginal revenue curve, the marginal cost curve, the average total cost curve, and the average variable cost curve faced by managers. To maximize profit, managers need to produce the output Q_M where the marginal cost curve intersects that of marginal revenue. If the monopolist produces Q_M , she will set a price of P_M . Because she is the only member of her market, her firm's demand curve is the industry demand curve. This is in contrast to perfect competition, where the demand curve for a firm's output is horizontal. The demand curve for the monopolist's output slopes downward to the right, as shown in Figure 8.4.

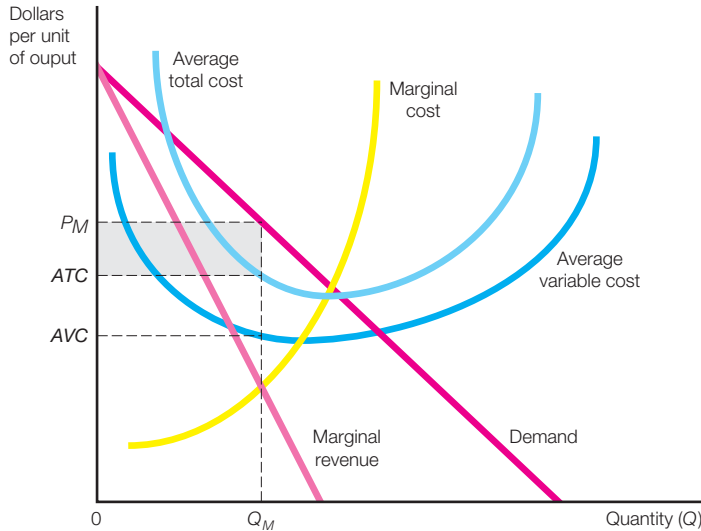
In Figure 8.4 managers generate profit per unit of $P_M - ATC$. This, multiplied by the number of units, Q_M , is the shaded area of the figure and equals total profit. Note also that $P_M > AVC$ to fulfill the second managerial rule of profit maximization.

Relative to managers in perfectly competitive markets, monopolists choose a higher price and lower output. This lets them charge a price higher than marginal cost and hence generate economic profit. Managers in perfectly competitive

FIGURE 8.4

Output and Price Decisions of a Monopolist

In equilibrium, the monopolist produces Q_M units of output and sets a price of P_M . (Note that, in contrast to perfect competition, the demand curve slopes downward to the right.)



markets can only set price equal to marginal cost. In the preceding example, if managers were forced to behave as a perfect competitor, they would set price equal to marginal cost; that is, $P = 10 - Q = 1 + Q = MC$. This yields $2Q = 9$ or an output of 4.5 and a price of 5.5; that is, $P = 10 - 4.5$. Therefore, output is curtailed under monopoly (from 4.5 to 3), price is increased (from \$5.5 to \$7), and profit is increased (from \$9.125 to \$12.5)—see Table 8.1.

To see that the monopolist's price exceeds marginal cost, recall that

$$MR = P[1 - (1/|\eta|)]$$

and that the monopolist sets marginal revenue equal to marginal cost. Therefore,

$$MC = P \left[1 - \left(\frac{1}{|\eta|} \right) \right] \text{ or} \quad (8.3)$$

$$P = \frac{MC}{\left[1 - \left(\frac{1}{|\eta|} \right) \right]}$$

Because $|\eta| > 1$, it follows that $[1 - (1/|\eta|)] < 1$, which means P must exceed MC .

STRATEGY SESSION: Franchiser versus Franchisee?

What happens when a franchiser with monopoly power has a different objective than a franchisee has? Consider the case of McDonald's, where the franchiser makes its money by collecting a percentage of each store's gross sales or total revenues (formally called a *royalty*). Therefore the franchiser wants to maximize the total revenue from the sales of its hamburgers by having each store maximize its total revenues (and by adding more stores). We assume that franchisees wish to maximize their profit. To do so, the franchisee would set $MR = MC$ (in the elastic range of demand where $|\eta| > 1$). But maximizing total revenue requires that $MR = 0$ (and that $|\eta| = 1$). Because the objectives of the franchiser and the franchisee cannot be accomplished with the same pricing policy, they are in conflict. The situation is depicted in the figure.^a

Note that this conflict exists in other situations such as book publishing. Authors generally receive a percentage of book sales as a royalty. Therefore, authors would like publishers to maximize total revenues. Publishers, on the other hand, want to maximize profit. Hence the author's objective would entail lower book prices than the publisher's; don't blame us for the high price of this book!

The royalty rates^b (as a percentage of monthly sales) for some popular franchises are shown here:

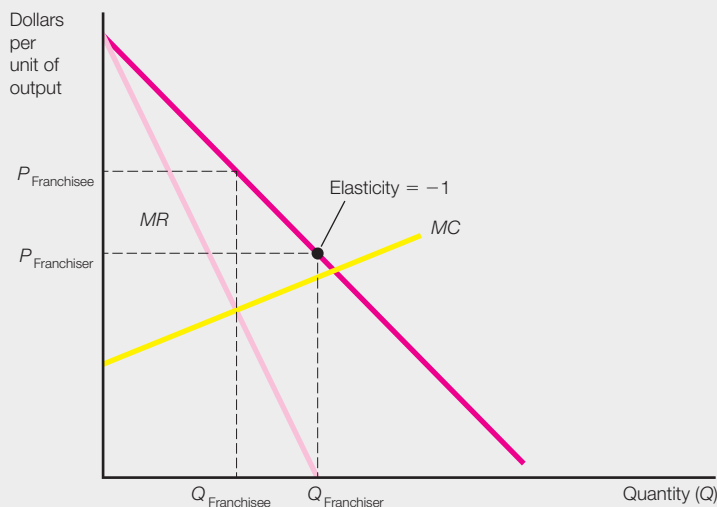
McDonald's	4% of gross sales
DQ Grill and Chill Restaurants	4% of gross sales
DQ Treat Centers	6% of gross sales
Motel 6	4% of gross room revenues
Studio 6	5% of gross room revenues
UPS Store	5% of gross sales and gross commissions
Ben and Jerry's	3% of gross sales
Jiffy Lube	5% of gross sales

In addition, there is generally a marketing or sales promotion fee (for advertising) of 2–6% of sales for most franchises as well as a one-time franchise fee. Papa John's, however, has a royalty fee of 5% of *net* sales (and hence the incentives of the franchiser and franchisees are aligned).

^a*Businessweek*, June 2, 1997.

^bCompany websites, May 2008.

A monopolist produces at $P_{\text{Franchisee}}, Q_{\text{Franchisee}}$ as dictated by the rule that marginal revenue equals marginal cost. A revenue maximizer produces where marginal revenue is zero ($P_{\text{Franchiser}}, Q_{\text{Franchiser}}$)—that is, where $|\eta| = 1$.



PROBLEM SOLVED: The McComb Company

To illustrate how managers choose price and output to maximize profit, consider the situation facing managers at the McComb Company, a monopolist producing and selling a product with the demand curve

$$P = 30 - 6Q \quad (8.4)$$

where P is price (in thousands of dollars) and Q is the firm's output (in thousands of units). The firm's total cost function is

$$TC = 14 + 3Q + 3Q^2 \quad (8.5)$$

where TC is total cost (in millions of dollars).

From the demand curve in equation (8.4), we determine the firm's total revenue (in millions of dollars), which is

$$TR = PQ = (30 - 6Q)Q = 30Q - 6Q^2$$

Therefore, marginal revenue equals $\Delta TR/\Delta Q = 30 - 12Q$.

From the total cost function in equation (8.5), we determine that marginal cost is

$$\Delta TC/\Delta Q = 3 + 6Q$$

Setting marginal revenue equal to marginal cost gives us

$$MR = 30 - 12Q = 3 + 6Q = MC$$

This means that $Q = 1.5$. Inserting 1.5 for Q in the demand equation (8.4), we find that $P = 30 - 6(1.5)$, or \$21. So to maximize profit, managers should set a price of \$21,000 and produce and sell 1,500 units. Doing so will result in profit equal to $[30(1.5) - 6(1.5)^2] - [14 + 3(1.5) + 3(1.5)^2] = \6.25 million.

The extra profit earned by monopoly managers is generated by their ability to choose a price greater than marginal cost, whereas the perfect competitor merely charges the marginal cost. Of course managers in both markets must choose a price that is higher than average variable cost.

COST-PLUS PRICING

Unfortunately academic studies of pricing behavior consistently suggest that many managers do not price optimally. Instead they use the simple heuristic of **cost-plus pricing** (see Strategy Session: Markup Pricing on the iPad2, page 268).

Many managers act as if cost is the primary driver of price. Although this simple strategy guarantees that price is higher than the estimated average cost, it does not necessarily optimize profit. The behavior has many variations but follows a guiding principle: Price is set as a function of cost. Managers first allocate unit costs conditional on a given output level (such as 70% capacity). Then they add a profit margin. This margin is generally a percentage of costs and is added to the estimated average costs. The markup is meant to cover costs that are difficult to allocate to specific products and as a return on firm investment.

Cost-plus pricing Simplistic strategy that guarantees that price is higher than the estimated average cost.

STRATEGY SESSION: Markup Pricing on the iPad2

Just like iPods and iPhones, the new iPads have taken the market by storm and are flying off the shelves and at a rate that rippled the robust debut sales of the first iPad.

Responding to strong consumer demand and competition from Motorola Mobility, Samsung, Research in Motion, Hewlett-Packard, and 59 other companies, iPad has captured the imagination of the market (bringing in many first-time buyers).

How does such monopoly power standing translate into a price markup over cost? According to IHS iSuppli (a research firm), materials for the iPad2 cost about \$326.60 for a version with 32 gigabytes of memory that works with the mobile phone standard

known as global system for mobile communications. This is a 2% increase over the costs of an earlier model.

The price of an iPad2 ranges from \$499 to \$829. Thus, the ratio of price to cost is somewhere between 1.53 and 2.54 creating a tidy variable profit (profit excluding fixed cost) for the “gadget of the year.”

Source: Adam Satariano, “Apple May Have Sold 500,000 iPad2 Tablets on Debut Weekend, Analyst Says,” *Bloomberg News*, March 14, 2011, at www.bloomberg.com/news/2011-03-14/apple-may-have-sold-500-000-ipad-2-tablets-analyst-estimates.html.

In basic algebra, the percentage markup of this strategy is expressed as

$$\text{Markup} = (\text{Price} - \text{Cost}) / \text{Cost} \quad (8.6)$$

Profit margin The price of a product minus its cost.

where the numerator (Price – Cost) is the **profit margin**. If the cost of a paperback book is \$4 and its price is \$6,

$$\text{Markup} = (6 - 4) / 4 = 0.50$$

or 50%. If we solve equation (8.6) for price, the result is

$$\text{Price} = \text{Cost}(1 + \text{Markup}) \quad (8.7)$$

which is the pricing formula described in the previous paragraph. In the case of the paperback book,

$$\text{Price} = 4(1 + 0.5) = \$6$$

because the markup is 50%.

Target return What managers hope to earn and what determines the markup.

Managers may also choose a **target return** they hope to earn, which determines the markup. For example, General Electric at times has established a target rate of return of 20%. Under a target rate of return pricing, price is set equal to

$$P = L + M + K + (F/Q) + (\pi A/Q) \quad (8.8)$$

where P is price, L is unit labor cost, M is unit material cost, K is unit marketing cost, F is total fixed or indirect costs, Q is the number of units managers plan to

produce during the relevant planning period, A is total gross operating assets, and π is the desired profit rate on those assets. If managers estimate unit labor cost at \$2, unit material cost at \$1, unit marketing cost at \$3, total fixed cost at \$10,000, output at 1,000 units, with assets of \$100,000 and a target rate of return of 15%, they should set price at

$$P = 2 + 1 + 3 + (10,000/1,000) + [0.15(100,000)/1,000] = \$31$$

One issue facing managers who produce more than one product is the charge for indirect cost, or overhead. Often managers use the heuristic of allocating this cost among the firm's products on the basis of their average variable costs. If a firm's total annual indirect costs (for all products) are estimated to be \$3 million and the total annual variable costs (for all products) are estimated to be \$2 million, indirect costs would be allocated to products at a rate of 150% of variable costs. For example, if the average variable cost of product Y is estimated to be \$10, managers should add a charge of $1.50 \times \$10$, or \$15, for indirect cost. Adding this charge to the average variable cost, the manager estimates the fully allocated cost of $\$10 + \15 , or \$25. Managers then set a price that is higher than this cost to generate profit. For example, if the markup is 40%, the price is $1.40 \times \$25$, or \$35.

COST-PLUS PRICING AT THERMA-STENT

Cost-plus pricing is widely used in medical group purchasing organizations. Therma-Stent is a producer of graft stents. Managers set price by estimating the average production costs (including indirect ones). They then add a 40% markup to set the product's market price

$$\begin{aligned}\text{Factory cost/unit} &= \$2,300 \text{ (at production of 20,000)} \\ 40\% \text{ markup} &= \$920 \\ \text{U.S. list price} &= \$3,220\end{aligned}$$

Using the heuristic eases the complexity of setting price by ignoring market considerations. For instance, price is set without considering prices of rival products. This pricing scheme works better when products are differentiated. Therma-Stent produces graft stents that are unique in form and surface structure.

COST-PLUS PRICING AT INTERNET COMPANIES AND GOVERNMENT-REGULATED INDUSTRIES

Managers at many online companies seem to have adopted a cost-plus pricing scheme. Consider OnSale, an online store. Managers have structured a pricing policy called "At Cost" where they sell products based at the wholesale price plus a fixed transaction fee (the markup).³

3. Henry Norr, "Egghead Whips Up a \$400 Million Deal with Onsale," *San Francisco Chronicle*, July 15, 1999.

STRATEGY SESSION: When Is the Price Too High?

Monopolists mark price up above marginal cost (which would be the price if the market were perfectly competitive), as we know from the text. They wouldn't (assuming rationality) charge a price higher than the one that would maximize their profit. But is that price/profit too high? From a social welfare perspective, we know that the answer is yes (because pricing at marginal cost maximizes social welfare). But we have many prices that exceed marginal costs in our markets, so when should we intervene and say that the markup is too high?

We have captive shipper provisions in the U.S. rail industry where the Surface Transportation Board can still regulate rates in what is basically a deregulated U.S. rail industry. Regulation can occur under a very specific set of conditions related to the overall revenue adequacy of the railroad, the ratio of the rate to the variable cost of the movement, and the cost of building a stand-alone railroad specifically to serve the captive shipper in question, among other things. Basically no rail carrier rates have been declared too high (although the shippers believe this to be the case, they can't win in a proceeding before the Surface Transportation Board).

But we aren't alone in this situation. Transnet Freight Rail in South Africa has recently raised its rates on a coal export railway line to Richards Bay by 26.3% over last year. The clients of the railroad have agreed to the increase. A lawyer has stated that since Transnet is the dominant rail carrier in South Africa, the rate is subject to the country's Competition Act, which states that the dominant group can't charge

excessively to the detriment of consumers. A case can be won, the lawyer claims, if it is shown that the rates were raised without the carrier experiencing a corresponding increase in costs.

This would seem to be an answer in search of a problem. The customers of the railroad have not filed a complaint. Transnet argues that costs have indeed increased. A director for Venmyn, a firm that advises mining companies on legal compliance with the law and project valuation, notes that if that's what it costs to put up and maintain that infrastructure, that's what one has to pay.

As in the United States, these cases are not brought before the competition authorities because they are "notoriously" difficult to prove as no one can determine what excessive pricing is. Only one case has been brought in South Africa, and it didn't lead to any conclusion and the parties settled.

So whereas laws in the United States and South Africa exist to prevent prices (rates) from becoming too high in the railroad industry (indicating that legislators were convinced that situations could exist where prices were too high), in practice, the implementation of these laws in both countries has been sparse (indicating that the potential plaintiffs in these cases don't believe that it is possible to prove such a case even though they believe the prices are too high).

Source: Allan Seccombe, "Transnet Freight Rail Tariffs Run into Criticism," *Business Day* (South Africa), April 5, 2011, at www.businessday.co.za/Articles/Content.aspx?id=139223.

Online sellers in other product lines have adopted the same pricing scheme. Managers at www.bybb.com sell major household appliances (GE and Hotpoint washers, dryers, refrigerators, and ovens) and electronics (Toshiba, Mitsubishi, and Sony TVs and MP3 players) on a cost-plus basis, where the purchaser is shown the wholesale price of the item.⁴

4. Cost Appliances and Electronics: Lowest Prices on the Web, at www.bybb.com.

Many automobile dealers also use a cost-plus pricing scheme, though they tend to make it difficult for consumers to accurately determine cost. Auto dealer invoices contain some items such as area allowances, which are hard for the novice consumer to interpret, and manufacturer givebacks are not included in the invoice. Therefore consumers do not see the true price the dealer paid for the car. In addition, many customers trade in their old vehicles. In haggling with a customer, the dealers are concerned with how much money they can make in the package—the trade-in plus the sale of the new vehicle. This makes buying from the auto dealer harder than buying from the appliance seller (where trade-ins are nonexistent).

Government regulators also use cost-plus pricing in industries they regulate or control. For instance, the Coal Ministry in India recently allowed Rajmahal coal to be priced at cost plus Rs143 per tonne after meeting production costs.⁵ The danger of such a pricing scheme in a government-controlled industry is that, when the profit is guaranteed, firm managers may lose the incentive to be cost efficient. This tends to create a larger government regulatory bureaucracy to monitor costs.

CAN COST-PLUS PRICING MAXIMIZE PROFIT?

The important question we need to consider is how good a heuristic cost-plus pricing is for managers to use. So far it seems unlikely that cost-plus pricing will often maximize profit. Indeed, this pricing technique seems simple-minded in that it does not explicitly account for important factors on both the demand and supply sides. It certainly does not explicitly consider the extent of demand or the product's price elasticity, including the pricing behavior of rivals. On the supply side it looks at average, not marginal, cost. Nevertheless, if applied well, cost-plus pricing may result in managers almost maximizing profit. The possibility that cost-plus pricing is sometimes a good heuristic revolves around what factors managers consider in determining the size of the percentage markup or the target rate of return. For example, why was the markup on the paperback book cited earlier 50%? Why not 25% or 150%?

In choosing a markup to maximize profit, managers must estimate the book's price elasticity of demand. To understand why this is true, recall from equation (8.3) that

$$MC = P \left[1 - \left(\frac{1}{|\eta|} \right) \right]$$

Dividing both sides of the equation by $1 - (1/|\eta|)$, we get

$$P = MC \left\{ \frac{1}{\left[1 - \left(\frac{1}{|\eta|} \right) \right]} \right\}$$

5. B. Sanyal, "Cost-Plus Pricing Helps Rajmahal Expansion," August 26, 1996, at www.hindubusinessline.com/1996/08/26/BLFP.08.html.

So in setting price, if managers want to maximize profit, they need to understand how marginal cost and price elasticity of demand are associated. Formally, managers need to set price so it equals the products marginal cost multiplied by

$$\left\{ \frac{1}{1 - \left(\frac{1}{|\eta|} \right)} \right\}$$

Intuitively, equation (8.7) says that in cost-plus pricing, managers choose a price equal to cost multiplied by $(1 + \text{Markup})$. If managers use marginal (not average) cost, then a markup of

$$\text{Markup} = |\eta| / (|\eta| - 1) \quad (8.9)$$

will maximize profit.

Put differently, a manager can maximize profit using cost-plus pricing with a markup equal to the value specified in equation (8.9). As equation (8.9) clearly shows, the optimal markup depends on the product's price elasticity of demand. To help managers think about this, we have constructed Table 8.3, which shows the profit-maximizing markup corresponding to elasticity values. For example, if a product's price elasticity of demand equals -1.2 , the optimal markup is 500%. If its price elasticity is -21 , the optimal markup is only 5%. Table 8.3 should be studied carefully because it provides useful information to help managers choose an effective pricing policy.

Note the negative association between elasticity and markup. As the price elasticity of demand decreases (in absolute value), the optimal markup increases. Table 8.3 shows this clearly. To see that the inverse relationship in Table 8.3 between markup and price elasticity is reasonable, ask yourself the following question: If the quantity demanded of a product is not very sensitive to its price, should I set

TABLE 8.3

Relationship between Optimal Markup and Price Elasticity of Demand

Price Elasticity of Demand	Optimal Percentage Markup of Marginal Cost
-1.2	500
-1.4	250
-1.8	125
-2.5	66.67
-5.0	25
-11.0	10
-21.0	5
-51.0	2

PROBLEM SOLVED: The Humphrey Corporation

To illustrate how managers can use cost-plus pricing and maximize profit, consider the Humphrey Corporation, a seller of office furniture. One of Humphrey's major products is a metal desk for which the company pays \$76 per desk, including transportation and related costs. Although managers at Humphrey face a variety of overhead and marketing costs, these costs are essentially fixed, so marginal cost is approximately \$76. Given that many firms in Humphrey's geographic area sell reasonably comparable desks, Humphrey's managers estimate that the demand for desks is fairly price elastic—about -2.5 . On the basis of Table 8.3, managers should choose a markup of 66.67% to maximize profit.

According to equation (8.7), the optimal price is

$$\begin{aligned}\text{Price} &= \text{Cost}(1 + \text{Markup}) \\ &= 76(1 + 0.6667) \\ &= \$126.67\end{aligned}$$

So if managers want to maximize profit, they should choose the price of \$127 per desk. In so doing, managers will approximate profit-maximizing behavior. Behavior is only approximated because values of factors like marginal cost and price elasticity of demand must be estimated. Recognizing that this is the case, Humphrey's managers should be prepared to slightly alter the price once they see the market reaction to it.

a relatively high or low price for this product? Obviously you should set a high price if you want to make as much money as possible. Since this is what Table 8.3 tells us, it accords with common sense.

THE MULTIPLE-PRODUCT FIRM: DEMAND INTERRELATIONSHIPS

Managers at monopolies that produce multiple products face a more complex decision. Managers need to recognize that a change in the price or quantity sold of one product may influence the demand for other products. For example, if the Akkina Company produces and sells two products, X and Y , its total revenue (that is, sales) is represented as

$$TR = TR_X + TR_Y \quad (8.10)$$

where TR_X is its total revenue from product X and TR_Y is the total revenue from product Y . The marginal revenue from product X is

$$MR_X = \frac{\Delta TR_X}{\Delta Q_X} + \frac{\Delta TR_Y}{\Delta Q_X} \quad (8.11a)$$

and the marginal revenue from product Y is

$$MR_Y = \frac{\Delta TR_Y}{\Delta Q_Y} + \frac{\Delta TR_X}{\Delta Q_Y} \quad (8.11b)$$

The last term in each of these equations represents the demand interrelationship between the two products. In equation (8.11a), the last term shows the effect of an increase in the quantity sold of product X on the total revenue from product Y . This effect can be positive or negative. If products X and Y are complements, this effect is positive because an increase in the quantity sold of one product increases the total revenue from the other product. On the other hand, if products X and Y are substitutes, this effect is negative: An increase in the quantity sold of one product reduces the total revenue of the other product.

Managers who do not understand or pay proper attention to demand interrelationships of this sort can make serious pricing mistakes. For example, if product X is a fairly close substitute for product Y and the division of the Akkina Company producing product X launches a campaign to increase its sales, the results may be good for the division but bad for the company as a whole. Why? Because the resulting increase in product X 's sales may occur largely at the expense of product Y 's sales.

PRICING OF JOINT PRODUCTS: FIXED PROPORTIONS

In addition to being interrelated on the demand side, some products also have interrelated production characteristics. For example, products sometimes are produced in a fixed ratio, as in the case of cattle, where beef and hide are obtained from each animal. In such a situation there is no reason to distinguish between the products from the point of view of production or costs; managers should think of them as a bundle. One hide and two sides of beef might be such a bundle in the case of cattle because they are produced from each animal. With such jointly produced products, there is no economically correct way to allocate the cost of producing each bundle to the individual products.

To determine the optimal price and output of each such bundled product, managers need to compare the marginal revenue generated by the bundle to its marginal cost of production. If the marginal revenue—that is, the sum of the

QUANT OPTION

The marginal revenue from product X is

$$MR_X = \frac{\partial TR_X}{\partial Q_X} - \frac{\partial TR_Y}{\partial Q_X}$$

and the marginal revenue from product Y is

$$MR_Y = \frac{\partial TR_Y}{\partial Q_Y} - \frac{\partial TR_X}{\partial Q_Y}$$

PROBLEM SOLVED: Pricing Filet Mignon at the Palm Restaurant

If you want a steak in New York City, a top-notch place to go is one of the four Palm restaurants in the city. They are owned by the Palm Management Corporation, which has 26 U.S. locations (and one in London), all with the same menu. The price of a 9-ounce filet mignon steak in each of these 26 U.S. restaurants in April 2011 was as follows:

New York (4 locations), Chicago, East Hampton, Las Vegas, Los Angeles (2 locations)	\$43
Atlantic City, Boston, Charlotte, Denver, Houston, Miami, Nashville, Orlando, San Diego, San Juan, Tyson's Corner, Washington, DC	\$42
Atlanta, Dallas, Philadelphia, San Antonio, Tampa	\$41

Problems

1. Suppose that the corporate management at Palm notes that people in various cities, such as San Antonio, are more price conscious than people in other cities, such as New York, and that local managers must know what price to charge to be competitive in their local market. Assuming that the market for restaurant food is monopolistically competitive in each of these cities, is the demand for filet mignon at the Palm restaurant in San Antonio the same as at one of the Palm restaurants in New York? If not, how does it differ?
2. Suppose corporate management also states that their labor costs in New York are higher than they pay in other cities, such as San Antonio. Is the marginal cost curve for a filet mignon at a Palm restaurant in New York the same as at the Palm restaurant in San Antonio? If not, how does it differ?
3. Why is the price of a filet mignon dinner higher in New York than in San Antonio?
4. If the marginal cost is 20% higher in New York than in San Antonio, and the price elasticity of demand is -3 in New York and -4 in San Antonio,

what would you expect to be the percentage price differential between New York and San Antonio?

Solutions

1. No. On the basis of corporate management's statement, the demand curve is more price elastic in San Antonio than in New York. A 1% price increase is likely to reduce the quantity demanded by a larger percentage in San Antonio than in New York.
2. No. The marginal cost curve is lower in San Antonio than in New York.
3. As pointed out on page 265, the profit-maximizing price equals

$$P = MC / [1 - (1/|\eta|)]$$

where MC equals marginal cost and $|\eta|$ equals the absolute value of the price elasticity of demand. (This is true under any market structure.) Because MC is higher and η is less elastic in New York than in San Antonio, the profit-maximizing price is higher in New York than in San Antonio.

4. If P_S is the price in San Antonio and P_N is the price in New York, MC_S is the marginal cost in San Antonio, MC_N is the marginal cost in New York, η_S is the price elasticity of demand in San Antonio, and η_N is the price elasticity of demand in New York, then

$$\begin{aligned} P_S/P_N &= \{MC_S/[1 - (1/|\eta_S|)]\} / \{MC_N/[1 - (1/|\eta_N|)]\} \\ &= \{MC_S/[1 - (1/|-4|)]\} / \{MC_N/[1 - (1/|-3|)]\} \\ &= \{MC_S/[1 - (1/4)]\} / \{MC_N/[1 - (1/3)]\} \\ &= \{MC_S/(3/4)\} / \{MC_N/(2/3)\} \\ &= \{(4/3)MC_S\} / \{(3/2)(1.2MC_S)\} \\ &= 1.33/1.8 = 0.74 \end{aligned}$$

Therefore, we would expect the price in San Antonio to be about 26% below the price in New York.

Source: <http://www.thepalm.com/files/files/AprilDinner>, accessed on July 28, 2011.

Note that the total marginal revenue curve coincides with the marginal revenue curve for product A at all outputs beyond Q_0 in Figure 8.5. This is true because managers should never sell product B at a level where its marginal revenue is negative. A negative marginal revenue means managers can increase revenue by selling fewer units. Therefore, if the total output exceeds Q_0 , managers should sell only part of the product B produced; specifically, they want to sell the amount corre-

Optimal Pricing for Joint Products Produced in Fixed Proportions (Case 1)

The graph illustrates the pricing and output decisions of a monopolist. The vertical axis represents the price in dollars per unit of output, and the horizontal axis represents the output of product bundles per period. The graph shows the following curves and points:

- Total marginal revenue:** A downward-sloping pink line starting from the vertical axis.
- Marginal cost:** An upward-sloping yellow curve.
- Marginal revenue for A:** A downward-sloping magenta line.
- Demand for product A:** A downward-sloping magenta line, parallel to and above the marginal revenue for A.
- Marginal revenue for B:** A downward-sloping blue line.
- Demand for product B:** A downward-sloping blue line, parallel to and above the marginal revenue for B.

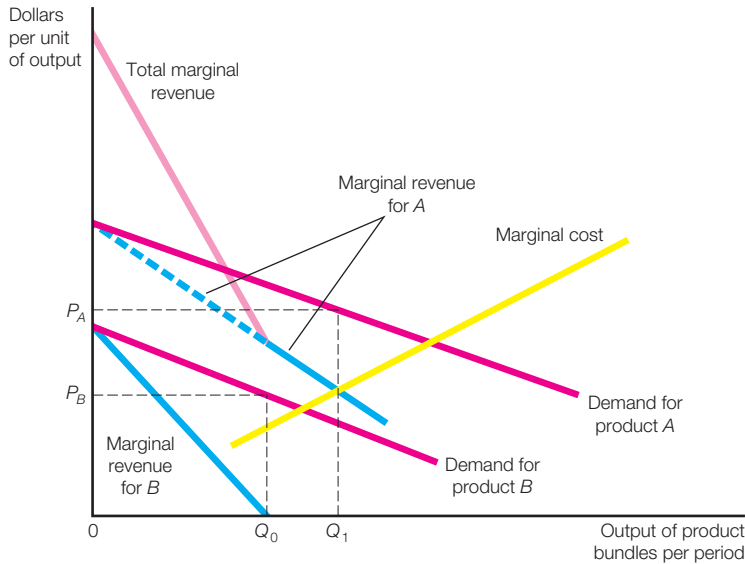
The profit-maximizing output Q is determined where Marginal Cost equals Marginal Revenue for A. This output level Q is projected down to the Demand for product A curve to find the price P_A . The price P_B is found on the Demand for product B curve at output Q . The output Q_0 is the point where the Demand for product B curve intersects the horizontal axis.

6. For simplicity, we assume that the demand curve for product *A* is not influenced by the price of product *B* and vice versa.

FIGURE 8.6

Optimal Pricing for Joint Products Produced in Fixed Proportions (Case 2)

The price of product A is set at P_A , the price of product B is set at P_B , and not all of product B is sold.



sponding to an output of Q_0 product bundles. Consequently, if output exceeds Q_0 , the total marginal revenue equals the marginal revenue of product A alone.

What should managers do if the marginal cost curve intersects the total marginal revenue curve to the right of Q_0 in Figure 8.5? In particular, suppose managers face the situation shown in Figure 8.6, where the marginal cost curve is lower than in Figure 8.5 (but the other curves are the same). The profit-maximizing output is Q_1 , where the marginal cost and total marginal revenue curves intersect. All of product A produced is sold, the price being P_A ; but not all of product B is sold. Instead the amount sold is limited to the amount of output Q_0 , so that the price of product B is P_B . The “surplus” amount of product B (that is, $Q_1 - Q_0$) must be kept off the market to avoid depressing its price.

OUTPUT OF JOINT PRODUCTS: VARIABLE PROPORTIONS

Having discussed the case in which two joint products are produced in fixed proportions, we turn to the case in which they are produced in variable proportions. This generally is a more realistic case, particularly if a manager is considering

PROBLEM SOLVED: Profit Maximizing at Humphrey

Humphrey managers now face the following situation. They manufacture two different conference table legs that are cut from the same piece of metal. They differ in their design but are jointly produced in equal quantities. That is, for every unit of the modern design produced, Humphrey also produces a unit of classical design. Managers face the total cost function

$$TC = 100 + Q + 2Q^2 \quad (8.12)$$

where Q is the number of units (in tens) of output. [Each unit contains one classic leg and one modern leg.] The demand curves for the firm's two products are

$$P_A = 200 - Q_A \quad (8.13)$$

$$P_B = 150 - 2Q_B \quad (8.14)$$

where P_A and Q_A are the price and output (in tens) of the modern leg and P_B and Q_B are the price and output (in tens) of the classic leg.

Humphrey managers need to know how many units of each leg they should produce to maximize profit. The firm's total revenue is equal to the total revenues from its two products

$$TR = P_A Q_A + P_B Q_B \quad (8.15)$$

Substituting the right sides of equations (8.13) and (8.14) for P_A and P_B , respectively, it follows that

$$\begin{aligned} TR &= (200 - Q_A)Q_A + (150 - 2Q_B)Q_B \\ &= 200Q_A - Q_A^2 + 150Q_B - 2Q_B^2 \end{aligned}$$

Assuming that Humphrey managers want to sell all they produce of both products, $Q_A = Q_B = Q$ because a unit of one product is produced whenever a unit of the other product is produced. Therefore,

$$\begin{aligned} TR &= 200Q - Q^2 + 150Q - 2Q^2 \\ &= 350Q - 3Q^2 \end{aligned} \quad (8.16)$$

To obtain total profit, π , managers must subtract the total cost in equation (8.12) from the total revenue in equation (8.16)

$$\begin{aligned} \pi &= (350Q - 3Q^2) - (100 + Q + 2Q^2) \\ &= -100 + 349Q - 5Q^2 \end{aligned}$$

To maximize profit, we need to set $\Delta\pi/\Delta Q = 0$

$$\Delta\pi/\Delta Q = 349 - 10Q = 0$$

or

$$10Q = 349$$

Isocost curve Curve showing the amounts of goods produced at the same total cost.

Isorevenue lines Lines showing the combinations of outputs of products that yield the same total revenue.

a fairly long period. Even cattle's proportions of hides and beef can be altered because the animals can be bred to produce more or less beef relative to hide.

Suppose a firm produces and sells two products, A and B , and that each **isocost curve** (labeled TC in Figure 8.7) shows the amounts of these goods produced at the same total cost. The isocost curve labeled $TC = 13$ shows the various combinations of outputs—for example, 26 units of product A and 10 units of product B or 8 units of product A and 30 units of product B —that can be produced at a total cost of \$13,000 per day.

Also included in Figure 8.7 are **isorevenue lines** (labeled TR), each of which shows the combinations of outputs of the two products that yield the same total revenue. For example, the isorevenue line labeled $TR = 52$ shows the various

or

$$Q = 34.9$$

In other words, to maximize profit, Humphrey managers should produce 34.9 (tens of) legs of each design. Equation (8.13) tells managers they need to choose a price of \$165.10 to sell 34.9 (tens of) modern legs

$$P_A = 200 - 34.9 = \$165.10$$

And equation (8.14) tells managers they need to price classic legs at

$$P_B = 150 - 2(34.9) = \$80.20$$

to sell 34.9 (tens of) of them.

Managers are not quite through with the analysis yet. As indicated, we assume that Humphrey sells all it produces of both products. To see whether this is true, we must see whether, if $Q = 34.9$, the marginal revenues of both products are nonnegative. Only then will Humphrey managers sell all that is produced of both products (recall Figure 8.6). From equations (8.13) and (8.14), we find that TR_A , the total revenue from product A, equals

$$TR_A = P_A Q_A = (200 - Q_A)Q_A = 200Q_A - Q_A^2$$

and that TR_B , the total revenue from product B, equals

$$TR_B = P_B Q_B = (150 - 2Q_B)Q_B = 150Q_B - 2Q_B^2$$

Hence the marginal revenue of product A is

$$MR_A = \Delta TR_A / \Delta Q_A = 200 - 2Q_A = 130.2 \quad (\text{when } Q_A = 34.9)$$

And the marginal revenue of product B is

$$MR_B = \Delta TR_B / \Delta Q_B = 150 - 4Q_B = 10.4 \quad (\text{when } Q_B = 34.9)$$

Because both marginal revenues (MR_A and MR_B) are nonnegative when Q_A and $Q_B = 34.9$, the assumption underlying the analysis is valid.^a

^a If one product's marginal revenue had been negative when Q_A and Q_B equaled 34.9, the optimal solution would have involved producing more of this product than is sold, as indicated in Figure 8.6. The firm would sell only the amount of this product where the marginal revenue is zero. The marginal revenue for the other product would be used to determine its optimal amount level, as shown in Figure 8.6.

QUANT OPTION

$$d\pi/dQ = 349 - 10Q = 0 \text{ so that } 10Q = 349 \text{ or } Q = 34.9$$

The marginal revenue for product A is

$$MR_A = dTR_A/dQ_A = 200 - 2Q_A = 130.2 \text{ (when } Q_A = 34.9)$$

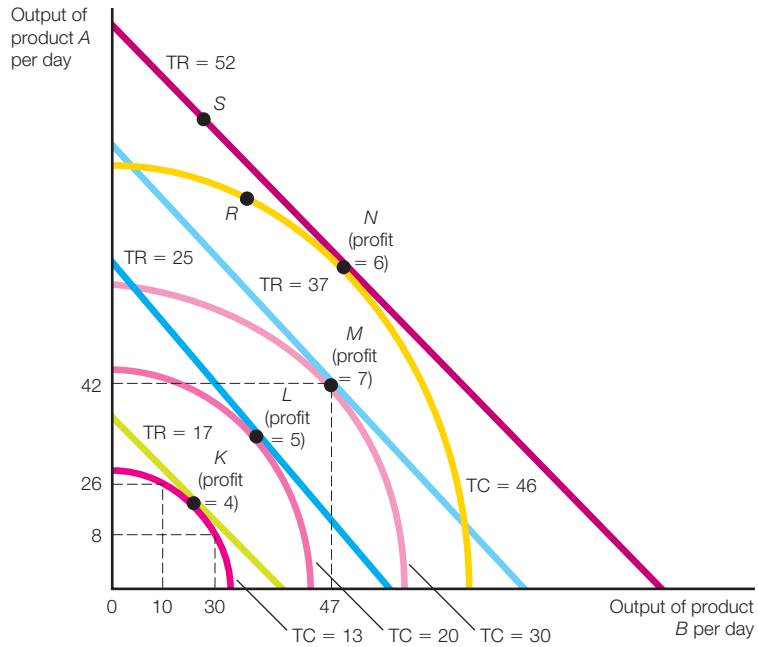
The marginal revenue for product B is

$$MR_B = dTR_B/dQ_B = 150 - 4Q_B = 10.4 \text{ (when } Q_B = 34.9)$$

FIGURE 8.7

Optimal Outputs for Joint Products Produced in Variable Proportions

The optimal point, which must be at a point where an isorevenue line is tangent to an isocost curve, is at point *M*, where profit per day is \$7,000.



combinations of outputs, such as those corresponding to points *S* or *N*, that yield a total revenue of \$52,000 per day. Other isorevenue lines show the output combinations that yield total revenues of \$17,000, \$25,000, and \$37,000, respectively.

The problem facing the manager is to determine how much of products *A* and *B* to produce. The first step toward solving this problem is to observe that if an output combination is at a point where an isorevenue line is *not* tangent to an isocost curve, it *cannot* be the optimal output combination. To see this, note that if an output combination is at a point where an isorevenue line is not tangent to an isocost curve (say point *R*), it is possible to increase revenue (without changing cost) by moving to a point (on the same isocost curve) where an isorevenue line is tangent to the isocost curve (say point *N*). Therefore, any output combination that is not at a tangency point cannot be the profit-maximizing output combination because we indicated how profit can be increased if the firm is at such a nontangency point.

Given that this is the case, we find the optimal output combination by comparing the profit level at each tangency point and choosing the point where the

profit level is the highest. For example, four tangency points are shown in Figure 8.7: points *K*, *L*, *M*, and *N*. As indicated in Figure 8.7, the profit levels corresponding to these four points are \$4,000, \$5,000, \$7,000, and \$6,000, respectively. So if we must choose among the output combinations on the isocost curves in Figure 8.7, the optimal output combination for this firm is point *M*, where the managers produce and sell 42 units of product *A* and 47 units of product *B* per day and make a profit of \$7,000.

MONOPSONY

Whereas monopoly occurs when there is one seller, **monopsony** occurs when there is one buyer. As in monopoly, the monopsonist controls price. Consider the market for busboys for New York City restaurants. There are many restaurants and many aspiring busboys. If a restaurant wants to hire an additional busboy, it can pay the prevailing wage for busboys, and that wage won't change as a result of its hiring. However, consider The Company in a company town. When it wishes to hire another worker, because it employs such a large proportion of the labor force, it will influence the wage.

The demand for labor is labor's marginal revenue product—that is, the incremental revenue that an additional worker will generate for The Company and the additional benefit of hiring another worker. Formally, it is The Company's marginal revenue multiplied by the marginal product of labor. It is downward-sloping because marginal revenue falls as output increases and because labor's marginal product falls as more labor is employed (recall the law of diminishing marginal productivity from Chapter 5). Denote the labor supply curve as $P = c + eQ$, where P is the wage of labor and Q is the number of workers willing to work at that wage. Note that it is upward-sloping, reflecting the influence that the monopsonist has on the prevailing wage rate; that is, to hire another worker, The Company must increase the wage to entice a worker either into the workforce or away from another job (and by so doing will have to pay all its workers the new higher wage).

The Company's total expenditure on labor (total cost) is

$$C = PQ = (c + eQ)Q = cQ + eQ^2$$

QUANT OPTION

The marginal cost of hiring another worker is

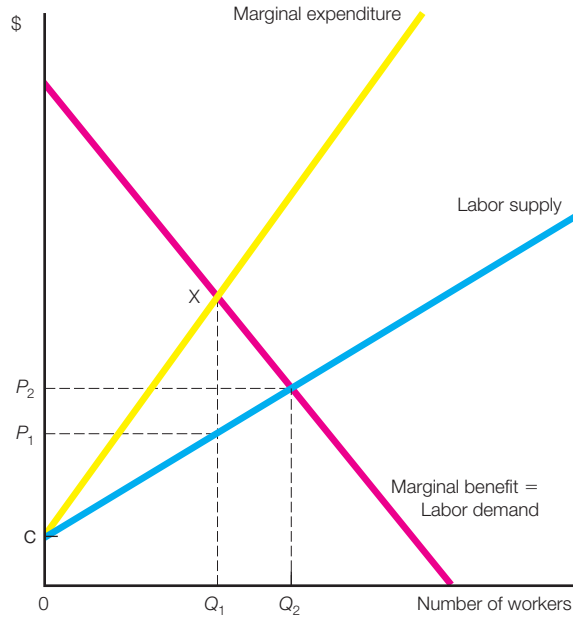
$$dC/dQ = c + 2eQ$$

Monopsony Markets that consist of a single buyer.

FIGURE 8.8

Optimal Monopsony Pricing

The optimal number of workers hired under monopsony is less (Q_1) than the optimal number of workers hired under perfect competition (Q_2); and the optimal wage under monopsony (P_1) is less than the optimal wage under perfect competition (P_2).



To maximize profit, managers will equate the marginal benefit of hiring another worker with the marginal expenditure (marginal cost) of hiring another worker

$$\Delta C / \Delta Q = c + 2eQ = MC$$

Figure 8.8 shows the optimal amount of labor and the wage paid by the monopsonist. The marginal benefit equals the marginal expenditure at point X with Q_1 workers. Dropping down to the labor supply curve yields a wage of P_1 . Note that the monopsonist restricts the amount of labor hired (Q_1) and pays a lower wage (P_1) than it would if the labor market were perfectly competitive (Q_2 and P_2).

MONOPOLISTIC COMPETITION

We now turn our attention to monopolistic competition. From a managerial point of view, a central characteristic of monopolistic competition is product differentiation. Unlike perfect competition, in which all managers sell an identical product,

firms in monopolistic competition sell similar but not identical products. Hence consumers can associate specific products with a given firm. For example, in retail markets both American Apparel and Gap sell similar tank tops for women. Managers at each firm make their tank tops slightly different (color, fabric, design). Managers could also offer different services or use different distribution channels to differentiate their products. Due to the differences among their products, managers have some control over product price, though price differentials are relatively small because the products of other firms are so similar.

In perfectly competitive markets, the firms included in an industry are easy to determine because they all produce an identical product. But when managers can differentiate their products, it is not as simple to accurately define an industry. Each firm produces a somewhat different product. Nevertheless, it is often useful to group together firms that produce similar products and call them a **product group**. We can define a product group called *neckties* or *toothbrushes* or *shirts*. The process by which we combine firms into product groups is somewhat arbitrary; there is no way to decide how close a pair of substitute products must be to belong to the same product group. Clearly the broader the definition of a product group, the greater the number of firms included.

Product group Group of firms that produce similar products.

In addition to product differentiation, other conditions must be met for an industry to qualify as one of monopolistic competition:

1. *There must be many firms in the product group.* The product must be produced by perhaps fifty to a hundred or more firms, with each firm's product a fairly close substitute for the products of the other firms in the product group.
2. *The number of firms in the product group must be large enough that each firm expects its actions to go unheeded by its rivals and unimpeded by possible retaliatory moves on their part.* Hence, when formulating their own price and output policies, they do not explicitly concern themselves with their rivals' responses. If there are many firms, this condition normally is met.
3. *Entry into the product group must be relatively easy, and there must be no collusion, such as price fixing or market sharing, among managers in the product group.* It generally is difficult, if not impossible, for a great many firms to collude.

Price and Output Decisions under Monopolistic Competition

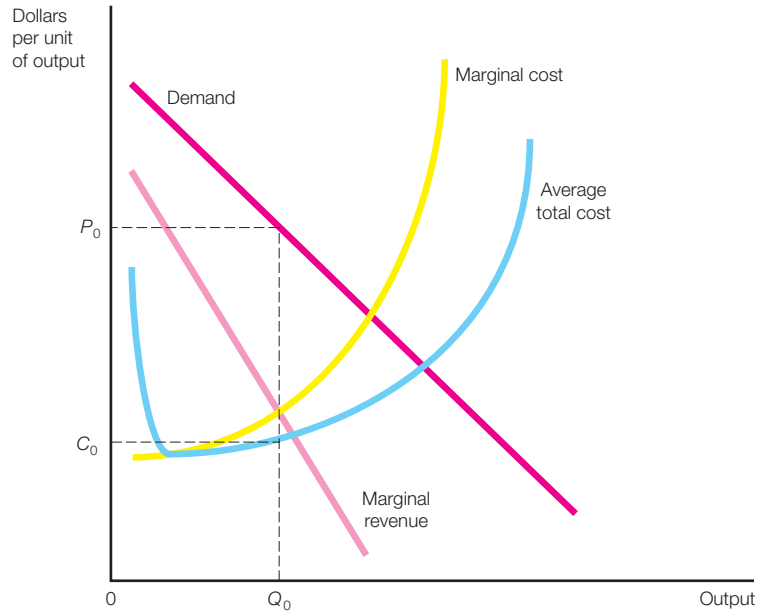
If each firm produces a somewhat different product, it follows that the demand curve facing each manager slopes downward to the right. That is, if the firm raises its price slightly, it will lose some, but by no means all, of its customers to other firms. And if it lowers its price slightly, it will gain some, but not all, of its competitors' customers.

Figure 8.9 shows the short-run equilibrium of a monopolistically competitive firm. Managers, in the short run, set price at P_0 and output at Q_0 because this combination of price and output maximizes profit. We can be sure this combination of

FIGURE 8.9

Short-Run Equilibrium in Monopolistic Competition

The firm will set its price at P_0 and its output rate at Q_0 because marginal cost equals marginal revenue at this output. It will earn a profit of $P_0 - C_0$ per unit of output.



price and output maximizes profit because marginal cost equals marginal revenue at this output level. Economic profit is earned because price, P_0 , exceeds average total cost, C_0 . As in the case of monopoly and perfect competition, price must exceed average variable cost for profit to be maximized (clearly this occurs in Figure 8.9—average variable cost lies below average total cost).

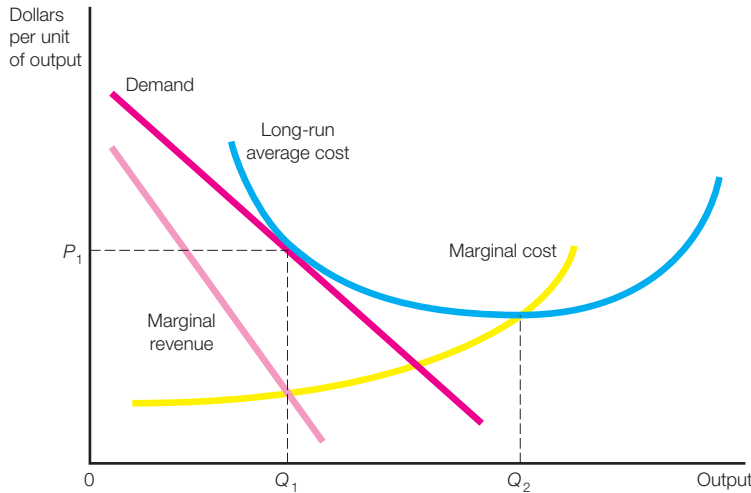
One condition for long-run equilibrium in these markets is that each firm makes no economic profit or loss because entry or exit of firms will occur otherwise—and entry and exit are incompatible with long-run equilibrium. Another condition for long-run equilibrium is that each firm maximize its profit. At what price and output are both these conditions fulfilled? Figure 8.10 shows that the long-run equilibrium is at a price of P_1 and an output of Q_1 . The zero economic profit condition is met at this combination of price and output because the firm's average cost at this output equals the price, P_1 . And the profit maximization condition is met because the marginal revenue curve intersects the marginal cost curve at Q_1 .⁷

7. The seminal work in the theory of monopolistic competition was E. Chamberlin, *The Theory of Monopolistic Competition* (Cambridge, MA: Harvard University Press, 1933).

FIGURE 8.10

Long-Run Equilibrium in Monopolistic Competition

The long-run equilibrium is at price P_1 and output Q_1 . There is zero profit because long-run average cost equals price. Profit is maximized because marginal cost equals marginal revenue at this output.



ADVERTISING EXPENDITURES: A SIMPLE RULE

Managers in monopolistic competition, as well as in other market structures, spend huge amounts on advertising. How much should a profit-maximizing manager spend on advertising? This section derives a simple rule to help managers answer this question.⁸ The quantity a firm sells of its product is assumed to be a function of its price and the level of its advertising expenditures. We assume diminishing marginal returns to advertising expenditures, which means that beyond some point, successive advertising outlays yield smaller increases in sales. (Table 8.4 shows an illustrative case in which successive increments of \$100,000 in advertising outlays result in smaller increases in quantity sold. For example, the quantity sold increases by 2 million units when advertising expenditures rise from \$800,000 to \$900,000, but by only 1.5 million units when they rise from \$900,000 to \$1 million.)

Let P be the price of a unit of the product and MC the marginal cost of production. If we assume that neither price nor marginal cost is altered by small changes in advertising expenditure, managers realize an increase in gross profit of $(P - MC)$ from each additional unit of product. Why is this the *gross* profit of selling an additional unit of output? Because it takes no account of whatever additional advertising expenditures are required to sell this extra unit of output. To

8. This rule, put forth by R. Dorfman and P. Steiner, applies to monopolistic or oligopolistic (see Chapter 11) firms as well as monopolistically competitive firms.

TABLE 8.4

Relationship between Advertising Expenditures and Quantity

Advertising Expenditures (Millions of Dollars)	Quantity Sold of Product (Millions of Units)
0.8	15.0
0.9	17.0
1.0	18.5
1.1	19.5
1.2	20.0

obtain the net profit, managers must deduct these additional advertising outlays from the gross profit.

To maximize net profit, a manager must set advertising expenditures at the level where an extra dollar of advertising results in extra gross profit equal to the extra dollar of advertising cost. Unless this is the case, a manager can increase the firm's total net profit by changing advertising outlays. If an extra dollar of advertising results in more than a dollar of increase in gross profit, the extra dollar should be spent on advertising (because this increases the total net profit). If the extra dollar (as well as the last dollar) of advertising results in less than a dollar's increase in gross profit, advertising outlays should be cut.⁹ Therefore, if ΔQ is the number of extra units of output sold as a result of an extra dollar of advertising, the manager should set advertising expenditures so that

$$\Delta Q(P - MC) = 1 \quad (8.17)$$

because the right side of this equation equals the extra dollar of advertising cost and the left side equals the extra gross profit resulting from this advertising dollar.

If we multiply both sides of equation (8.17) by $P/(P - MC)$, we obtain

$$P\Delta Q = P/(P - MC) \quad (8.18)$$

Because the manager is maximizing profit, he is producing an output level at which marginal cost (MC) equals marginal revenue (MR). Therefore, we can substitute MR for MC in equation (8.18), the result being

$$P\Delta Q = P/(P - MR) \quad (8.19)$$

Using equation (8.2), we can show that the right side of equation (8.19) equals $|\eta|$, the negative of the price elasticity of demand for the firm's product.¹⁰ The left side of equation (8.19) is the marginal revenue from an extra dollar of advertising (it equals the price times the extra number of units sold as a result of

9. For simplicity, we assume that the gross profit resulting from an extra dollar spent on advertising is essentially equal to the gross profit resulting from the last dollar spent. This is an innocuous assumption.

10. Recall from equation (8.2) that $MR = P[1 - (1/|\eta|)]$. Therefore, $[1 - (1/|\eta|)] = MR/P$ and $1/|\eta| = 1 - (MR/P)$; this means that $|\eta| = 1/[1 - (MR/P)] = P/(P - MR)$ which is the right side of equation (8.19).

an extra dollar of advertising). To maximize profit, the manager should set advertising expenditure so that

$$\text{Marginal revenue from an extra dollar of advertising} = |\eta| \quad (8.20)$$

This rule can be very helpful to managers.¹¹ Consider managers at the Humphrey Corporation, who estimate the price elasticity of demand for its product equals -1.6 . To maximize profit, managers must set the marginal revenue from an extra dollar of advertising equal to 1.6 , according to the rule in equation (8.20). Suppose managers believe an extra \$100,000 of advertising will increase sales by \$200,000. This association implies that the marginal revenue from an extra dollar of advertising is about \$200,000/\$100,000, or 2.0 rather than 1.6 . Because the marginal revenue exceeds the absolute value of the price elasticity, managers can increase profit by advertising more.¹² To maximize profit, managers should increase advertising to the point where the marginal revenue from an extra dollar of advertising falls to 1.6 —that is, the absolute value of the price elasticity of demand.

USING GRAPHS TO HELP DETERMINE ADVERTISING EXPENDITURE

A simple graphical technique can help managers identify optimal advertising expenditures. Take the case of the Hertzfeld Chemical Company. Curve A in Figure 8.11 shows the relationship between a product's price elasticity of demand and the amount managers spend on advertising. Managers should think of price elasticity as a proxy for the effectiveness of their differentiation strategies. Advertising is a strategic variable managers use to convey their differentiating message. The graph shows a curvilinear association between advertising and price elasticity. With little or no advertising, differentiation is slight between rival products; hence the price elasticity is high (in absolute value).

But because effective advertising can induce consumers to attach economic value to product attributes, increases in advertising spending reduce the product's price elasticity (in absolute value) considerably (by decreasing the product's perceived substitutability with other goods).¹³ For any advertising level, the B curve shows the marginal revenue from an extra dollar of advertising. Because the A curve intersects the B curve when Hertzfeld's advertising expenditure is R dollars, this, on the basis of equation (8.20), is the estimated optimal advertising spending.

Clearly the optimal advertising level depends on the position and shape of the B and A curves. For example, suppose Hertzfeld's B curve shifts rightward to B' , as shown in Figure 8.11. Such a shift might occur if managers increase advertising effectiveness. Because advertising is more effective, marginal revenue increases, so managers want to increase their ad spending. Hence we see the optimal level of advertising increases (to S dollars in Figure 8.11).

11. However, this rule is based on many simplifying assumptions and is not a complete solution to this complex problem.

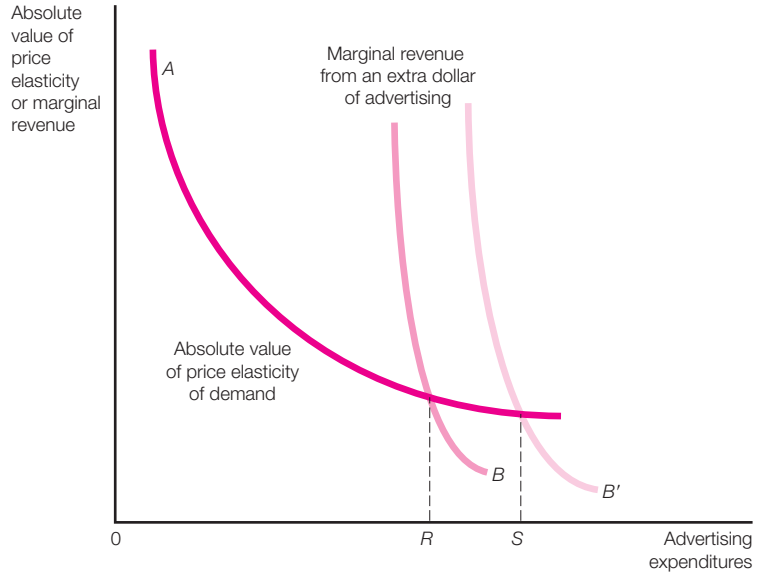
12. Had Humphrey's managers believed that the marginal revenue from an extra dollar of advertising was less than the price elasticity of demand, a *reduction* in the firm's advertising expenditures would increase profit.

13. This is true for some products, but not for others. In some cases the absolute value of price elasticity of demand is directly, not inversely, related to the amount spent on advertising.

FIGURE 8.11

Optimal Advertising Expenditure

The firm's optimal advertising expenditure is R if the marginal revenue curve is B (or S if the marginal revenue curve is B').



ADVERTISING, PRICE ELASTICITY, AND BRAND EQUITY: EVIDENCE ON MANAGERIAL BEHAVIOR

Promotions and advertising tend to be two sides of the same coin. Although they both seek to improve market performance, promotions appeal to the price-sensitive, whereas ads build brand loyalty. Promotions use a price-oriented message to test the limits of brand loyalty; advertising illuminates brand worth and does not mention price. Both strategies persuade consumers by influencing the price sensitivities of consumers. Promotions increase price elasticity and, in the long run, limit the price consumers are willing to pay for brand quality. Understanding these effects on consumer behavior certainly helps managers better understand the consequences of their actions.

So can we find real-world evidence of the effectiveness of these strategies? Yes. A large body of evidence shows that promotions increase the price elasticities of consumers.¹⁴ These studies also show that the change in elasticities varies across consumers and time. In addition, brand loyalty does protect against promotions: Promotional strategies have less effect on the elasticities of brand loyalists relative to nonloyalists. And promotion is characterized as decaying in time. Short-term

14. K. Pauwells, D. Hanssens, and S. Siddarth, "The Long-Term Effects of Price Promotions on Category Incidence, Brand Choice, and Purchase Quantity," *Journal of Marketing Research*, vol. 39 (November 2002), pp. 421–436; C. Mela, S. Gupta, and D. Lehmann, "The Long-Term Impact of Promotion and Advertising on Consumer Brand Choice," *Journal of Marketing Research*, vol. 34 (May 1997), pp. 248–262.

change is greater, as if consumers operate with a high discount rate; or they may have short memories.

Mela, Gupta, and Lehmann report on a mature good market in which the ratio of advertising to promotions shifted from spending \$250 million on advertising and offering promotions less than 10% of the time to spending less than \$100 million on advertising and giving discounts more than 25% of the time. They found that the price elasticity of the average nonloyalist was twice that of the average brand loyalist. Loyalty is measured as frequency of repeat purchases. A drop in advertising messages affects all consumers, but the effect is much greater on the nonloyal crowd. In fact, a drop in advertising leads to a larger number of non-loyalists. Without reinforcement, a brand is eroded by price. Frequent promotions encourage nonloyalists to look for them, so their price sensitivities increase. The effect of promotions on the price sensitivity of nonloyalists is four times that of loyalists.

Pauwells, Hanssens, and Siddarth report on the soup and yogurt markets. They studied buying habits over a two-year period by analyzing purchases of over 690,000 ounces of yogurt and 535,000 ounces of soup. In both markets they find the effects of a promotion on price sensitivities was the greatest within the two-week period following its announcement. After this initial period, the effect grew weaker. The frequency of promotions also varied across firms, as did the amount of the discounts. These findings indicated that promotions are used more frequently by managers whose brand loyalty is weaker.

SUMMARY

1. Under monopoly, a manager maximizes profit by setting output at the point where marginal revenue equals marginal cost. It does not follow that managers in monopoly markets always earn significant profit. If the monopolist cannot cover its variable costs, it, like a perfectly competitive firm, will shut down, even in the short run.

2. An industry that is monopolized generally sets a higher price and a lower output than if it were perfectly competitive. The perfectly competitive firm operates at a point where price equals marginal cost, whereas the monopolist operates at a point at which marginal revenue equals marginal cost (and price exceeds marginal cost).

3. Empirical studies indicate that cost-plus pricing is used by many managers. In this approach, a manager estimates the cost per unit of output (based on some assumed output level) and adds a markup to include costs that cannot be allocated to any specific product and to provide a return on the firm's investment. On the surface, it is questionable whether this approach can maximize profit; but if marginal cost (not average cost) is really what is being marked up and the size of

the markup is determined (in the appropriate way) by the product's price elasticity of demand, cost-plus pricing can approximate profit maximization.

4. Firms generally sell more than one product. It is important for managers to recognize the demand interrelationships among the products they sell. Also, products are often interrelated in production. If two products are produced jointly in fixed proportions, the profit-maximizing output occurs where the total marginal revenue curve (the vertical summation of the marginal revenue curves for the individual products) intersects the marginal cost curve for the bundle of products, assuming the marginal revenue of each product is nonnegative.

5. If two products are produced jointly in variable proportions, we can construct isocost curves, each of which shows the combinations of outputs that can be produced at the same total cost. Also, isorevenue lines can be constructed, each of which shows the combination of outputs that yield the same total revenue. For an output combination to be optimal, it must be at a point where an isorevenue line is tangent to an isocost curve. To determine which output combination is optimal, we compare the profit levels at the tangency points. The tangency point where profit is the highest is the optimal output combination.

6. Monopsony occurs when there is only one buyer. Analogous to monopoly, the monopsonist restricts the amount purchased to less than what would occur if perfect competition existed and decreases the price paid relative to the price that would prevail under perfect competition.

7. In contrast to perfect competition, where all firms sell an identical product, firms under monopolistic competition sell somewhat different products. Producers differentiate their products from those of other producers. Therefore, the demand curve facing each firm slopes downward to the right—and is not horizontal, as it would be under perfect competition. Each firm sets marginal revenue equal to marginal cost if it maximizes profit.

8. Managers of monopolistically competitive firms spend large amounts on advertising. To maximize its profit, a manager should set an advertising level so the marginal revenue from an extra dollar of advertising equals the absolute value of the price elasticity of demand (under the conditions discussed).

9. Advertising of price changes may increase the price elasticity of demand for the product whose price has changed. This happens because the advertising makes consumers more aware of the price changes. Measures of brand loyalty are useful in guiding decisions concerning promotional activities to increase sales of particular brands.

PROBLEMS

1. Harry Smith owns a metal-producing firm that is an unregulated monopoly. After considerable experimentation and research, he finds that the firm's marginal cost curve can be approximated by a straight line, $MC = 60 + 2Q$,

where MC is marginal cost (in dollars) and Q is output. The demand curve for the product is $P = 100 - Q$, where P is the product price (in dollars) and Q is output.

- a. If Smith wants to maximize profit, what output should he choose?
 - b. What price should he charge?
2. The Wilson Company's marketing manager has determined that the price elasticity of demand for its product equals -2.2 . According to studies she carried out, the relationship between the amount spent by the firm on advertising and its sales is as follows:

Advertising Expenditure	Sales
\$100,000	\$1.0 million
200,000	1.3 million
300,000	1.5 million
400,000	1.6 million

- a. If the Wilson Company spends \$200,000 on advertising, what is the marginal revenue from an extra dollar of advertising?
 - b. Is \$200,000 the optimal amount for the firm to spend on advertising?
 - c. If \$200,000 is not the optimal amount, would you recommend that the firm spend more or less on advertising?
3. The Coolidge Corporation is the only producer of a particular type of laser. The demand curve for its product is

$$Q = 8,300 - 2.1P$$

and its total cost function is

$$TC = 2,200 + 480Q + 20Q^2$$

where P is price (in dollars), TC is total cost (in dollars), and Q is monthly output.

- a. Derive an expression for the firm's marginal revenue curve.
 - b. To maximize profit, how many lasers should the firm produce and sell per month?
 - c. If this number were produced and sold, what would be the firm's monthly profit?
4. The Madison Corporation, a monopolist, receives a report from a consulting firm concluding that the demand function for its product is

$$Q = 78 - 1.1P + 2.3Y + 0.9A$$

where Q is the number of units sold, P is the price of its product (in dollars), Y is per capita income (in thousands of dollars), and A is the firm's advertising

expenditure (in thousands of dollars). The firm's average variable cost function is

$$AVC = 42 - 8Q + 1.5Q^2$$

where AVC is average variable cost (in dollars).

- a. Can we determine the firm's marginal cost curve?
 - b. Can we determine the firm's marginal revenue curve?
 - c. If per capita income is \$4,000 and advertising expenditure is \$200,000, can we determine the price and output where marginal revenue equals marginal cost? If so, what are they?
5. The Wilcox Company has two plants with the marginal cost functions¹⁵

$$MC_1 = 20 + 2Q_1$$

$$MC_2 = 10 + 5Q_2$$

where MC_1 is marginal cost in the first plant, MC_2 is marginal cost in the second plant, Q_1 is output in the first plant, and Q_2 is output in the second plant.

- a. If the Wilcox Company minimizes its costs and produces five units of output in the first plant, how many units of output does it produce in the second plant? Explain.
 - b. What is the marginal cost function for the firm as a whole?
 - c. Can we determine from these data the average cost function for each plant? Why or why not?
6. If the Rhine Company ignores the possibility that other firms may enter its market, it should set a price of \$10,000 for its product, which is a power tool. But if it does so, other firms will begin to enter the market. During the next two years it will earn \$4 million per year, but in the following two years it will earn \$1 million per year. On the other hand, if it sets a price of \$7,000, it will earn \$2.5 million in each of the next four years because no entrants will appear.
- a. If the interest rate is 10%, should the Rhine Company set a price of \$7,000 or \$10,000? Why? (Consider only the next four years.)
 - b. If the interest rate is 8%, should the Rhine Company set a price of \$7,000 or \$10,000? Why? (Consider only the next four years.)
 - c. The results in parts (a) and (b) pertain to only the next four years. How can the firm's managers extend the planning horizon?
7. During recessions and economic hard times, many people—particularly those who have difficulty getting bank loans—turn to pawnshops to raise cash. But even during boom years, pawnshops can be profitable. Because the collateral that customers put up (such as jewelry, guns, or electric guitars) is generally worth at least double what is lent, it generally can be sold at a profit. And because usury laws allow higher interest ceilings for pawnshops than for other lending institutions, pawnshops often charge spectacularly high rates of inter-

15. This question pertains to the chapter appendix.

est. For example, Florida's pawnshops charge interest rates of 20% or more *per month*. According to Steven Kent, an analyst at Goldman, Sachs, pawnshops make 20% gross profit on defaulted loans and 205% interest on loans repaid.

- a. In 2012 there were about 15,000 pawnshops in the United States. This was much higher than in 2007, when the number was about 12,000. Why did the number increase?
 - b. In a particular small city, do the pawnshops constitute a perfectly competitive industry? If not, what is the market structure of the industry?
 - c. Are there considerable barriers to entry in the pawnshop industry? (*Note:* A pawnshop can be opened for less than \$250,000, but a number of states have tightened licensing requirements for pawnshops.)
8. In 1996 dairy farmers, hurt by a decade of low milk prices, began reducing their herds. Subsequently Kenneth Hein, a Wisconsin farmer, said he was getting \$16 per 100 pounds of milk, rather than \$12, which he had gotten earlier.¹⁶
- a. Why did the price increase?
 - b. Dairy cattle are often fed corn. When Hein got \$16 per 100 pounds of milk, he paid \$5 a bushel for corn; but when he got \$12 per 100 pounds of milk, he paid \$2.50 a bushel for corn. Does this mean that Hein made less money when the price of milk was \$16 than when it was \$12?
9. The demand for diamonds is given by

$$P_Z = 980 - 2Q_Z$$

where Q_Z is the number of diamonds demanded if the price is P_Z per diamond. The total cost (TC_Z) of the De Beers Company (a monopolist) is given by

$$TC_Z = 100 + 50Q_Z + 0.5Q_Z^2$$

where Q_Z is the number of diamonds produced and put on the market by the De Beers Company. Suppose the government could force De Beers to behave as if it were a perfect competitor—that is, via regulation, force the firm to price diamonds at marginal cost.

- a. What is social welfare when De Beers acts as a single-price monopolist?
 - b. What is social welfare when De Beers acts as a perfect competitor?
 - c. How much does social welfare increase when De Beers moves from monopoly to competition?
10. The Hassman Company produces two joint products, X and Y . The isocost curve corresponding to a total cost of \$500,000 is

$$Q_Y = 1,000 - 10Q_X - 5Q_X^2$$

where Q_Y is the quantity of product Y produced by the firm and Q_X is the quantity of product X produced. The price of product X is 50 times that of product Y .

16. *Philadelphia Inquirer*, September 14, 1996.

- a. If the optimal output combination lies on this isocost curve, what is the optimal output of product X?
 - b. What is the optimal output of product Y?
 - c. Can you be sure that the optimal output combination lies on this isocost curve? Why or why not?
11. The McDermott Company estimates its average total cost to be \$10 per unit of output when it produces 10,000 units, which it regards as 80% of capacity. Its goal is to earn 20% on its total investment, which is \$250,000.
 - a. If the company uses cost-plus pricing, what price should it set?
 - b. Can it be sure of selling 10,000 units if it sets this price?
 - c. What are the arguments for and against a pricing policy of this sort?
 12. The Morrison Company produces tennis rackets, the marginal cost of a racket being \$20. Because there are many substitutes for the firm's rackets, the price elasticity of demand for its rackets equals about -2 . In the relevant range of output, average variable cost is very close to marginal cost.
 - a. The president of the Morrison Company feels that cost-plus pricing is appropriate for his firm. He marks up average variable cost by 100% to set price. Comment on this procedure.
 - b. Because of heightened competition, the price elasticity of demand for the firm's rackets increases to -3 . The president continues to use the same cost-plus pricing formula. Comment on its adequacy.
 13. The Backus Corporation makes two products, X and Y. For every unit of good X that the firm produces, it produces two units of good Y. Backus's total cost function is

$$TC = 500 + 3Q + 9Q^2$$

where Q is the number of units of output (where each unit contains one unit of good X and two units of good Y) and TC is total cost (in dollars). The demand curves for the firm's two products are

$$\begin{aligned} P_X &= 400 - Q_X \\ P_Y &= 300 - 3Q_Y \end{aligned}$$

where P_X and Q_X are the price and output of product X and P_Y and Q_Y are the price and output of product Y.

- a. How much of each product should the Backus Corporation produce and sell per period?
- b. What price should it charge for each product?

EXCEL EXERCISE: SIMPLE MONOPOLY

Suppose that the monopolist has the following estimate for her demand curve:

Price (P)	Quantity Demanded (Q)
14	0
13	1
12	2
11	3
10	4
9	5
8	6
7	7
6	8
5	9
4	10
3	11
2	12
1	13
0	14

and has the following estimate of her variable cost of producing each output:

Quantity Produced	Variable Cost
0	0
1	2.5
2	6.0
3	10.5
4	16.0
5	22.5
6	30.0
7	38.5
8	48.0
9	58.5
10	70.0
11	82.5
12	96.0
13	110.5
14	126.0

The firm has fixed costs of 5.

From the first spreadsheet exercise in Chapter 1, we know the demand equation is $Q = 14 - P$, which can also be expressed as $P = 14 - Q$ (by adding P to both sides of the equation and subtracting Q from both sides of the equation). As shown in the text, the firm's marginal revenue curve given this demand curve is $MR = 14 - 2Q$ —that is, same intercept, double the slope.

Determining the relationship between quantity produced and variable cost is a bit tougher, but it can be shown that every number in the variable cost column comes from the formula $VC = 2Q + 0.5Q^2$. We know from the text that marginal cost $MC = dVC/dQ = 2 + Q$ and that to maximize profit, the firm sets $MR = MC$. Doing so yields

$$MR = 14 - 2Q = 2 + Q = MC$$

or $3Q = 12$

or $Q = 4$

Substituting $Q = 4$ into the demand curve gives $P = 14 - 4 = 10$.

The firm's total revenue (TR) is $TR = P \cdot Q = 10 \cdot 4 = 40$.

The firm's variable cost (VC) is $VC = 2Q + 0.5Q^2 = 2 \cdot 4 + 0.5 \cdot 4 \cdot 4 = 8 + 8 = 16$.

The firm's fixed cost (FC) = 5.

The firm's profit (π) = $\pi = TR - TC - FC = 40 - 16 - 5 = 19$.

But suppose you didn't have the above equations and did not know the calculus. Could you still calculate what the profit-maximizing output and price are and how much profit you will make?

The answer is yes, and with the spreadsheet, complete enumeration is quite simple.

Open up the spreadsheet and enter 14 in cell A1, then 13 in cell A2, and so on, until you've sequentially ordered all the prices in the first table above.

Then enter 0 in cell B1, 1 in cell B2, and so on, until you've entered all quantities demanded from the table above opposite their respective prices.

Then enter $=A1 \cdot B1$ in cell C1. This is the firm's total revenue if it sells at price 14. You only have to enter the formula in cell C1. You then click on the lower right-hand corner of cell C1 with your mouse and drag down to cell C15. That will transform the correct total revenue formula to each cell. You can then see from column C how total revenue first rises, reaches a maximum, and then falls as price is decreased as described in the text.

Then enter $=C2 - C1$ in cell D2. This shows the change in total revenue as you go from making no sales at $P = 14$ to making one sale at $P = 13$ —that is, your marginal revenue from making the first sale. As with the total revenue, you don't have to add the formula each time; just use the same click-and-drag method described for total revenue. This marginal revenue column enables you to see how marginal revenue decreases as sales increase and how it ultimately becomes negative (which, as we show on page 262, is a place no profit-maximizing manager

wants to go). We call this the *discrete marginal revenue* because it assumes that one can only sell products in integers (think of buying automobiles or cans of beans).

In our formula version above ($MR = 14 - 2Q$), we assume that output can be purchased on a continuous basis (think of your purchases of gasoline or ham at the deli). For reference, enter $=14-2*B1$ in cell E1, and so on, to compare the continuous marginal revenue with the discrete marginal revenue. As above, you only have to enter the one formula in cell E1, and then click and drag.

You will notice that the discrete marginal revenue is different (and greater) than the continuous marginal revenue. This reflects the difference between integer changes in output and very small changes in output. Notice that where total revenue is maximized, the continuous marginal revenue is 0, as shown on page 266.

Enter 5 in cell F1, and so on. This is the firm's fixed cost (FC). Use the click-and-drag method to fill out the column.

Enter 0 in cell G1, 2.5 in cell G2, 6 in cell G3, and so on; in other words, enter the numbers from the variable cost (VC) table above.

Enter $=F1+G1$ in cell H1 and so on via click and drag after the first entry. Column H is the firm's total cost—that is, $TC = FC + VC$.

Enter $=H2-H1$ in cell I2, and so on. Fill the column via click and drag. This column measures the firm's discrete marginal cost, or how total cost changes as one increases output by an integer (think of the production of automobiles).

The formula used above ($MC = 2 + Q$) to solve for the profit-maximizing output was the continuous marginal cost, assuming that you could produce output in very small increments (think gasoline production). In column J, we present the continuous marginal cost to compare with the discrete marginal cost of column I. Enter $=2+B1$ in cell J1, and so on, via click and drag. Notice that the discrete and continuous marginal costs differ reflecting the increments of output assumed and that the continuous marginal cost exceeds the discrete marginal cost.

Column K is our objective. It is the profit column, total revenue minus total cost. Enter $=C1-H1$ in cell K1, and so on, via click and drag.

Now search column K for the highest number. If you entered everything correctly, you should find 19 in profit and, reading leftward, find in row 5 that the price is 10 and the quantity is 4. You may also enter $=\text{Max}(K1:K15)$ in cell K16 to find the highest profit.

Thus without any calculus, you were able to find the profit-maximizing price and quantity. Note that at $P = 10$, the continuous marginal revenue equaled the continuous marginal cost ($=6$). Note also that at $P = 10$, the discrete marginal revenue equaled 7, while the discrete marginal cost equaled 5.5. Because cash flow in (7) exceeds cash flow out (5.5), the producer who can only produce in integers wants to produce the fourth unit. Note that at $Q = 5$, the discrete marginal revenue is 5 and the discrete marginal cost is 6.5. Because cash flow in (5) is exceeded by cash flow out (6.5), the producer who can only produce in integers will not want to produce the fifth unit. Thus, regardless of whether we use the discrete or

continuous numbers, we come to the same conclusion: Profit is maximized when four units are produced and sold at price 10.

We can also see from the table that monopolists mark price up over marginal cost (10 versus either 6 or 5.5).

Column L calculates the firm's price elasticity of demand. Recall the formula from the text is $\eta_D = (P/Q)(\Delta Q/\Delta P)$. From the first table $\Delta Q/\Delta P = -1$, and so $\eta_D = -P/Q$. Enter $-A2/B2$ in cell L2 and fill the column via click and drag. Notice the profit-maximizing firm produces in the elastic range of the demand curve (as we showed on page 262), and the elastic range of the demand curve exists where the marginal revenue is positive (and the inelastic range of the demand curve is where the marginal revenue is negative). You can also observe how total revenue changes when prices change in the elastic and inelastic ranges of the demand curve (as shown on page 263).

APPENDIX: ALLOCATION OF OUTPUT AMONG PLANTS

Many firms own and operate more than one plant. In this appendix we show how managers should allocate output among various plants. This is an important decision, and our results have major direct practical value. We consider the case of the Johnson Company, a monopolist; but our results are valid for any manager who exercises market power.

The Johnson Company, a monopolist that makes a particular type of fixture, operates two plants with marginal cost curves shown in columns 2 and 3 of Table 8.5, output being shown in column 1. Clearly, if the managers decide to produce only one unit of output per hour, they should use plant I because the marginal cost between zero and one unit of output is lower in plant I than in plant II. Hence for the firm as a whole, the marginal cost between zero and one unit of output is \$10 (the marginal cost between zero and one unit for plant I). Similarly, if the managers decide to produce two units of output per hour, both should be produced in plant I, and the marginal cost between the first and second units of output for the firm as a whole is \$12 (the marginal cost between the first and second units in plant I). If managers decide to produce three units of output per hour, two should be produced in plant I and one in plant II, and the marginal cost between the second and third units of output for the firm as a whole is \$14 (the marginal cost between zero and one unit of output for plant II). Alternatively, all three could be produced at plant I (the marginal cost between the second and third units of output in plant I is also \$14).

Going on in this way, we can derive the marginal cost curve for the firm as a whole, shown in column 4 of Table 8.5. To maximize profit, the manager should find the output level at which the marginal revenue equals the marginal cost of the firm as a whole. This is the profit-maximizing output level. In this case it is three or four units per hour. Suppose managers choose to produce four units.¹⁷ To

17. The firm is indifferent between producing three or four units. If it produces four, its total revenue is \$92,000 (23×4) and its variable cost is \$50,000 ($10 + 12 + 14 + 14$), yielding a variable-cost profit of \$42,000. If it produces three, the total revenue is \$78,000 (26×3) and the variable cost is \$36,000 ($10 + 12 + 14$), yielding a variable-cost profit of \$42,000. Because both plants already exist, their fixed costs must be paid and therefore are irrelevant in the short run. In the long run (if demand were predicted to remain constant), the firm could divest itself of plant II.

TABLE 8.5

Costs of the Johnson Company

Output per Hour	Marginal Cost ^a		Marginal Cost for Firm ^a (Dollars)	Price (Dollars)	Marginal Revenue ^a (Dollars)
	Plant I (Dollars)	Plant II (Dollars)			
1	10	14	10	40	—
2	12	18	12	30	20
3	14	22	14	26	18
4	20	26	14	23	14
5	24	30	18	20.8	12

^aThese figures pertain to the interval between the indicated output and one unit less than the indicated output.

determine what price to charge, they must estimate what price corresponds to this output on the demand curve. In this case, the answer is \$23.

At this point we have solved most of the Johnson Company's problems, but not quite all. Given that managers will produce four units of output per hour, how should they divide this production between the two plants? The answer is that they need to set the marginal cost in plant I equal to the marginal cost in plant II. Table 8.5 shows this means that plant I would produce three units per hour and plant II would produce one unit per hour. The common value of the marginal costs of the two plants is the marginal cost of the firm as a whole; this common value must be set equal to the marginal revenue if the firm maximizes profit.

Many managers use this technique to allocate output among plants. For example, electric power companies have developed computer programs to facilitate the job of allocating electricity demand (or "load") among plants in accord with this theoretical rule. These programs allow a central dispatcher, who is in constant communication with the plants, to compute quickly the optimal allocation among plants. The result has been millions of dollars of savings.

As a further illustration, consider the Chou Company, which has plants at Altoona, Pennsylvania, and at High Point, North Carolina. The total cost function for the Altoona plant is

$$TC_A = 5 + 9Q_A + Q_A^2$$

where TC_A is the daily total cost (in thousands of dollars) at this plant and Q_A is its output (in units per day). The total cost curve for the High Point plant is

$$TC_H = 4 + 10Q_H + Q_H^2$$

where TC_H is the daily total cost (in thousands of dollars) at this plant, and Q_H is its output (in units per day).

QUANT OPTION

The Chou Company's marginal revenue is

$$dTR/dQ = 31 - 2Q$$

The Chou Company's demand curve is

$$P = 31 - Q$$

and its total revenue is

$$TR = PQ = (31 - Q)Q = 31Q - Q^2$$

Therefore the Chou Company's marginal revenue curve is

$$MR = \Delta TR / \Delta Q = 31 - 2Q$$

Note that $Q = Q_A + Q_H$, P is price, and MR is the marginal revenue (in thousands of dollars per unit).

To maximize profit, managers must choose a price and output such that

$$MC_A = MC_H = MR \quad (8.21)$$

where MC_A is the marginal cost (in thousands of dollars) at the Altoona plant and MC_H is the marginal cost (in thousands of dollars) at the High Point plant.

The Altoona plant's marginal cost is

$$MC_A = \Delta TC_A / \Delta Q_A = 9 + 2Q_A$$

The High Point plant's marginal cost is

$$MC_H = \Delta TC_H / \Delta Q_H = 10 + Q_H$$

QUANT OPTION

The Altoona plant's marginal cost is

$$dTC_A/dQ_A = 9 + 2Q_A$$

and the High Point plant's marginal cost is

$$dTC_H/dQ_H = 10 + Q_H$$

According to equation (8.21), MC_A must equal MC_H . Therefore,

$$9 + 2Q_A = 10 + Q_H$$

or

$$Q_H = -1 + 2Q_A$$

Also, because equation (8.21) states that MC_A must equal MR ,

$$\begin{aligned} 9 + 2Q_A &= 31 - 2(Q_A + Q_H) \\ &= 31 - 2(Q_A - 1 + 2Q_A) \\ &= 33 - 6Q_A \end{aligned}$$

or

$$8Q_A = 24$$

Consequently, $Q_A = 3$. And because $Q_H = -1 + 2Q_A$, it follows that $Q_H = 5$. Moreover, $P = 23$ because $P = 31 - (Q_A + Q_H)$. In a nutshell, managers should charge \$23,000 per unit and produce three units per day at the Altoona plant and five units per day at the High Point plant.

PART 5

SOPHISTICATED MARKET PRICING

CHAPTER 9

LEARNING OBJECTIVES

Motivation for Price
Discrimination

Price Discrimination

Using Coupons and Rebates for
Price Discrimination

Peak Load Pricing

Two-Part Tariffs

Summary

Problems

Excel Exercise: Perfect Price
Discrimination

Excel Exercise: Third-Degree
Price Discrimination

Appendix: Two-Part Tariff with
Intersecting Demands

MANAGERIAL USE OF PRICE DISCRIMINATION

Price discrimination is common across many markets and products. A car dealership aims to sell each vehicle for the highest price it can get (as long as that price is higher than its reservation price). Airline managers segment their markets and sell the same seats at significantly different prices depending on when the tickets are purchased, whether the tickets are refundable, penalties for changing flight plans, and the like. College administrators use a price discrimination policy by awarding financial aid; there is a wide variety of aid packages within the student population.

In general, managers try to identify submarkets on the basis of an individual's price elasticity of demand. A car dealership is an example of first-degree price discrimination, where the dealer attempts to extract the reservation price of each buyer. In effect each buyer is a submarket. The airlines use one general pricing model to divide the market into at least two submarkets: a relatively price-insensitive business class and a relatively price-sensitive leisure class. Airlines are an example of third-degree price discrimination, where each airline tries to extract the average reservation price of those similar in price sensitivity. The airline pricing model is compromised somewhat by Internet firms such as Expedia, which search airline databases for the lowest fares. This gives the consumer more information about the range of fares available and enables the consumer to potentially get a lower fare than that available from dealing directly with the carrier.

MOTIVATION FOR PRICE DISCRIMINATION

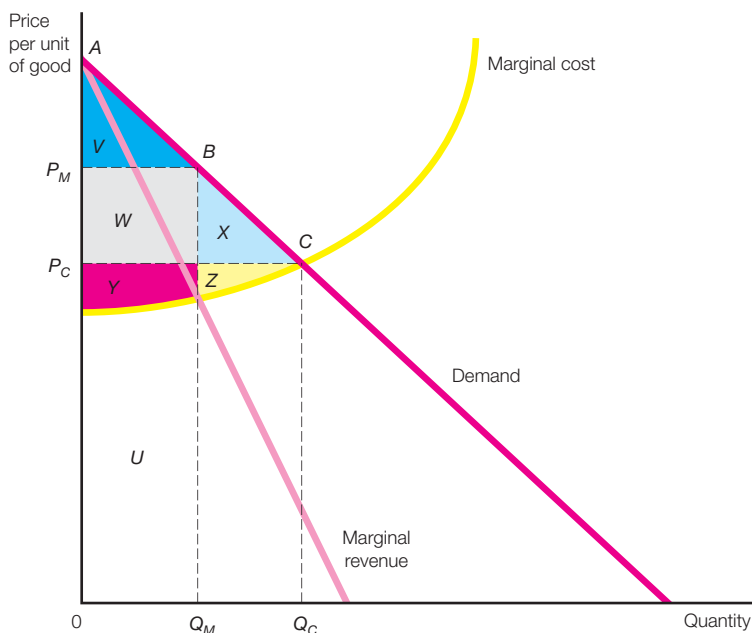
Consider Figure 9.1, which shows the profit-maximizing price and quantity for a single-price monopolist. By charging price P_M , the monopolist sells Q_M units. But aside from the customer whose reservation price was P_M , all other purchasing customers in area AB of the demand curve value the good at a price higher than P_M , but they are asked to pay only P_M for it. Consumers can retain a significant amount of consumer surplus—money they are willing to pay the producer but are not asked to do so. The amount of that consumer surplus is V (the area shaded in dark blue).

Consumers in area BC of the demand curve are unwilling to spend P_M for the good but have reservation prices that exceed the marginal cost of producing the good and hence represent potential profitable sales. These sales are not made by

FIGURE 9.1

Single-Price Monopolist Profit-Maximizing Outcome

The single-price monopolist prices at P_M and produces and sells Q_M units. Consumers in region AB are willing to pay a higher price than P_M yet are not asked to do so. Consumers in region BC are unwilling to pay a price as high as P_M but will pay a price higher than it costs the producer to make the good. Both these situations are potentially profitable sales that are not made.



the single-price monopolist, who curtails output at Q_M , whereas profitable sales could continue up to Q_C . The amount of profit represented by those potential sales is $X + Z$.

Instead the single-price monopolist settles for a variable-cost profit of $W + Y$: the gray plus red areas where total revenue is $P_M Q_M = W + Y + U$ and variable cost is the area U , under the marginal cost curve (as shown in Chapter 6).

If the monopolist raises the price above P_M to capture some of the consumer surplus in area V , area $X + Z$ becomes greater. If the manager lowers the price below P_M to capture some of the potential profit in area $X + Z$, area V becomes bigger. We know that managers cannot increase profit by deviating from P_M because it is the profit-maximizing price for the single-price monopolist. If the manager is going to capture some (or all) of region V and some (or all) of region $X + Z$, she cannot do it with a single-price strategy. Managers can capture surplus from area V and profit from areas $X + Z$ only with a strategy that involves two or more prices. We now explore what those strategies should be. Their motivation is capturing the additional profit in area V and areas $X + Z$. If the benefit of capturing that profit exceeds the costs of doing so (remember that sophisticated pricing is more costly to implement than simple single pricing), then our manager should do so.

PRICE DISCRIMINATION

Price discrimination When the same product is sold at more than one price.

Price discrimination occurs when the same product is sold at more than one price. For example, an airline may sell tickets on a particular flight at a higher price to business travelers than to college students. An automobile dealer may sell the exact same equipped make and model at different prices on the same day to different buyers. Even if the products are not precisely the same, price discrimination is said to occur if similar products are sold at prices that are in different ratios to their marginal costs. If managers sell boxes of candy with a label (costing \$0.02) saying “Premium Quality” in rich neighborhoods for \$12 and sell the same boxes of candy without this label in poor neighborhoods for \$5, this is price discrimination. Differences in price among similar products are not evidence of discrimination unless these differences do not reflect cost differences.

First-Degree Price Discrimination

Managers need to master three basic types of price discrimination: first, second, and third degree. The auto dealer is an example of the first degree, whereas the airline and candy firm are examples of the third degree. Selling electricity to certain customers is an example of the second degree. By examining price discrimination in a bit more detail, managers can better understand how to use all three types.

Consider again the diagram of a simple monopoly (single-price) profit maximizer shown in Figure 9.1. To reiterate, consumers in segment AB of the demand

STRATEGY SESSION: When Can You Haggle?

In general, Americans hate to haggle. They do so at car dealers (and detest the experience), in real estate transactions, in dealings with contractors for home repairs, in contract negotiations, and at flea markets; but most of their purchases involve posted prices, and most Americans accept that. However, haggling is the norm in many places around the world, and more haggling is drifting into the United States. Why is that?

According to a *Consumer Reports* National Research Center survey of over 2,000 shoppers, 61% bargained for products such as cell phones, furniture, medical bills, home electronics, household appliances, jewelry, antiques, and the like during the previous three years. In the home furnishings category, 94% reported paying less than the posted price. Of those who negotiated successfully, 61% reported savings of between \$50 and \$99; 26% reported savings of between \$1 and \$49; and 14% reported savings of \$100 or more.

Our MBA students play a series of managerial economics games against each other and MBA students at INSEAD's campuses in Fontainebleau and Singapore. Because they are French, INSEAD Fontainebleau gave a bottle of fine French champagne to the team that scored the highest. We decided to do the same. We teach 12 sections and hence needed 12 bottles. One of our faculty members went to several New Jersey liquor stores and negotiated a price for the champagne that was significantly below the posted price.

Here are *Consumer Reports'* tips for hagglers:

1. Use the power of timing. For service contracts, negotiate for discounts and perks at the time of the initial contract or at its renewal.
2. Offer cash. Credit and debit card companies charge merchants 2–8% for card use.
3. Look for flaws—scratches and dents that don't impair the performance of the product and can be hidden or covered by the purchaser.
4. Buy multiple units and ask for a quantity discount.

This opportunity to haggle is good news for consumers who don't buy goods at their posted prices because they are above their reservation prices. But it's got to be a good deal for the sellers too. Otherwise why would sellers negotiate?

However, it's potentially not good news for people who don't mind posted prices. If buying a bar stool becomes more like buying a car, many people are going to hate shopping. Even if you accept the posted price of the stool (which you know is high relative to the seller's reservation price given your car-buying experience), you'll worry that you paid too much; and you'll worry that your neighbor will quiz you about what you paid and then humble you when she tells you what she paid. At least when you both pay the same posted price, you feel equal to her.

Source: "Haggle Even at Stores; Survey Shows It Works," *Philadelphia Inquirer*, May 18, 2008, p. M-2.

curve are willing to pay more than the single monopoly price of P_M . Consumers in segment BC of the demand curve are willing to pay more for the good than it costs the producer to produce it—that is, the firm's marginal cost.

The simple monopolist makes a variable-cost profit of $W + Y$, as shown in Figure 9.1, and leaves the consumer surplus of V with the consumers of segment AB . If managers could perfectly price discriminate (another term for first-degree

PROBLEM SOLVED: Honest Sanjay's Use of First-Degree Price Discrimination

We now view an example of first-degree price discrimination versus simple monopoly pricing. Honest Sanjay sells used cars. The market demand for Sanjay's used cars is $P = 12 - Q$, where P is the price in thousands and Q is the quantity of cars sold per month.

Sanjay has two strategies of selling cars. He can set a price and merely pay a general manager to write the paperwork. The total cost of selling each car under such an arrangement is \$2 (thousand), so Sanjay's marginal cost is \$2 (thousand). This is also Sanjay's average variable cost of selling a car. Sanjay faces fixed costs of \$5 (thousand) per month.

To maximize profit under simple monopoly pricing, Sanjay should set marginal revenue equal to marginal cost. Sanjay's total revenue is $TR = PQ = (12 - Q)Q = 12Q - Q^2$. Sanjay's marginal revenue, $MR = \Delta TR / \Delta Q$, is

$$MR = 12 - 2Q$$

Setting Sanjay's marginal revenue equal to his marginal cost,

$$MR = 12 - 2Q = 2 = MC$$

gives $Q = 5$, which implies that the price of cars is $P = 12 - 5 = 7$ or \$7,000.

Sanjay's total revenue per month is \$35 (that is, $PQ = (\$7)(5) = \$35,000$), variable costs are \$10 (that is, $(AVC)Q = (\$2)(5) = \$10,000$), and fixed costs are \$5 (thousand), resulting in a monthly profit of \$20 (thousand) from simple monopoly pricing.

Sanjay could also sell cars the more common way—customers haggling with salespeople. Sanjay can hire a slick salesforce. By chatting with customers, a salesperson can pretty well estimate a customer's reservation price of a car; for example, salespeople often come right out and ask a customer how much the customer is looking to spend or are more subtle by asking, "What do you do for a living? Do you want to drive the car home tonight?" Salespeople who are not good at estimating customers' reservation prices tend not to be employed for long in the automobile business. A general manager is still needed to write the paperwork, and the salespeople are paid strictly on commission: \$1 (thousand) for each car they sell. Under this model of sales, Sanjay's marginal cost is \$3 (thousand) per car. The hagggle model is first-degree price discrimination in action. As mentioned, practicing price discrimination does carry costs not incurred by managers charging a single price.

Under this model, Sanjay sells cars up to the point where the reservation price equals marginal cost

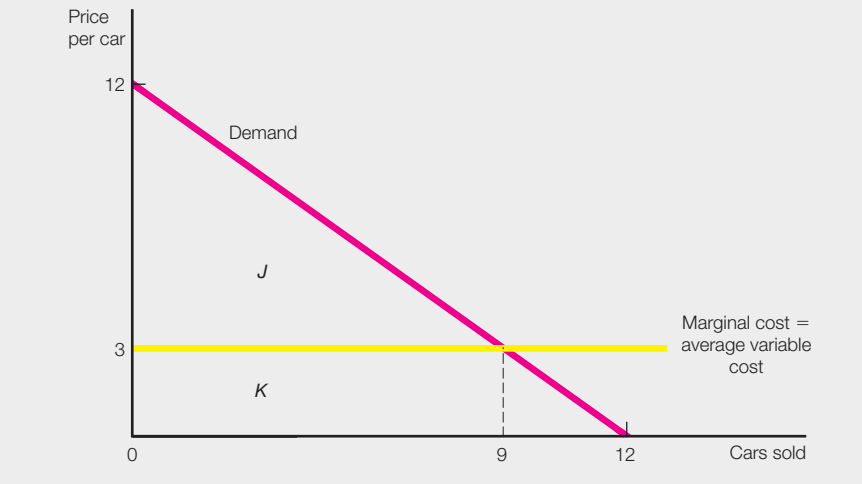
$$P = 12 - Q = 3 = MC \text{ or } Q = 9$$

price discrimination), they would charge the consumers in segment AB their reservation prices, capturing all the consumer surplus and turning it into producer surplus (that is, variable-cost profit). Note that when managers can perfectly discriminate in segment AB , the firm's variable-cost profit increases to $V + W + Y$.

First-degree discrimination lets managers expand sales. Because managers are not constrained by a single price, they can serve consumers in segment BC . This increases variable-cost profit by $X + Z$ because the reservation price of the consumers in segment BC exceeds the additional cost of producing the units

First-Degree Price Discrimination

The first-degree price discriminator captures all consumer surplus J and turns it into producer surplus (variable-cost profit).



All consumer surplus (J in the figure shown) is captured. Sanjay's profit is total revenue ($J + K$) less total cost (variable cost, K , plus fixed cost). In this case total revenue is \$67.5^a (thousand), variable cost is \$27 ($(AVC)(Q) = (\$3)(9) = \$27,000$), and fixed cost is \$5 (thousand), resulting in a profit of \$35.5 (thousand).

Sanjay prefers the haggle model over the simple monopoly posted-price model (because \$35.5 > \$20).

Presumably this explains why most auto dealers have not switched to the posted-price model and continue haggling to sell cars.

^aThe area of trapezoid $J + K$ is one-half the height (9) times the sum of the trapezoid's two sides (12 + 3). Therefore $0.5(9)(15) = 67.5$.

QUANT OPTION

Sanjay says, for those in the know, that marginal revenue is

$$dTR/dQ = 12 - 2Q$$

involved: $Q_C - Q_M$. By perfectly discriminating in both the AB and the BC segments, managers increase the firm's variable-cost profit (and hence its profit) by $V + X + Z$. This is precisely all the area we saw that the simple monopolist was not exploiting in Figure 9.1. The potential for this additional profit gets creative managers thinking about pricing strategies to capture it.

If managers can capture all of $V + X + Z$, we say they are practicing discrimination of the first degree. Managers always want to find ways to use first-degree price discrimination. In essence, the strategy allows them to charge each consumer his or her reservation price. By so doing, they guarantee that consumer surplus is zero. Clearly managers are willing to do this up to Q_C units in Figure 9.1. The additional revenue managers generate by selling an additional unit of product is the reservation price of the consumer. Managers sell to a consumer as long as the reservation price (which the manager can charge and the consumer is willing to pay) exceeds the marginal cost of production. In essence, in perfect discrimination the firm's demand curve becomes the firm's marginal revenue curve. Therefore, managers will not sell more than Q_C items because the marginal cost of producing them exceeds the revenue they will generate for the firm—that is, their reservation price.

Thus the profit-maximizing rule developed in Chapters 7 and 8 holds. The perfectly discriminating manager maximizes profit by producing until marginal revenue (represented by the demand curve) is equal to the output's marginal cost.

One interesting outcome of first-degree price discrimination is that it produces the same output as if the monopolist were in a perfectly competitive market—that is, Q_C . The difference between the two scenarios is in the distribution of consumer and producer surplus. In essence, using first-degree price discrimination, the manager gets to bake the cake and eat it too. In Figure 9.1, under perfectly competitive pricing (P_C), consumer surplus is $V + W + X$ and producer surplus is $Y + Z$. Because total welfare is the sum of consumer and producer surplus, social welfare is V through Z . Under first-degree discrimination, consumer surplus is zero (it has all been captured) and producer surplus is V through Z . Therefore, the welfare is the same under both pricing mechanisms, V through Z , but consumers benefit under perfect competition and producers get *all* the benefit of first-degree price discrimination. Because the output is the same in each pricing scheme, social welfare is identical.

For first-degree price discrimination, managers usually must have a relatively small number of buyers and must be able to estimate the maximum prices they are willing to accept. In addition, other conditions must hold that are elaborated on when we discuss third-degree price discrimination. For these reasons, the two-part tariff method of pricing (discussed later) is a simpler way to operationalize first-degree price discrimination in many situations.

The general retail market in the United States is not well suited to first-degree price discrimination. The market is predominantly posted price, so there is no haggling (with the exception of car buying, home buying, dealing with housing con-

tractors, and yard sales). In other cultures haggling is more prevalent. For example, in the bazaars of Asia, buyers are expected to haggle with sellers. In the finest pearl establishments of Hyderabad, India (a pearl capital of the world), transactions are all done by haggling (unless an unsuspecting tourist from a no-haggle country is uninformed enough to pay the price listed for the pearls). Negotiation is much more prevalent in business-to-business transactions in the United States.

Second-Degree Price Discrimination

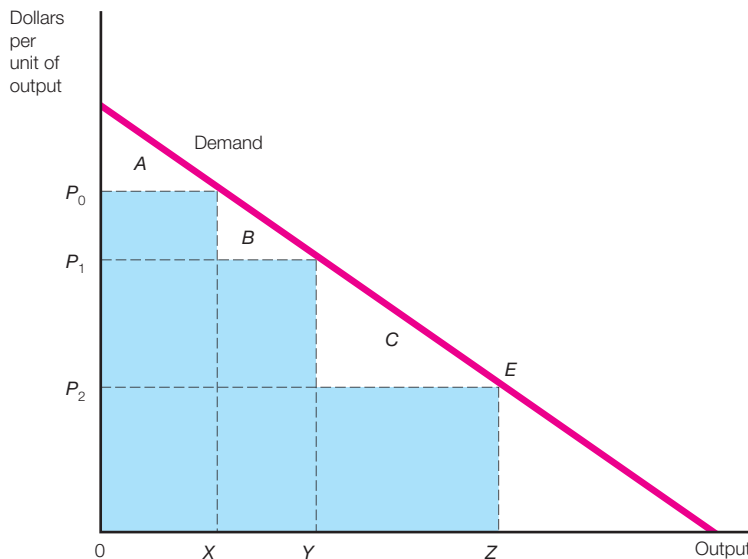
Second-degree price discrimination is most common in utility pricing. According to some authorities, second-degree price discrimination plays an important role in the schedule of rates charged by many public utilities—gas, water, electricity, and others.¹

Consider a gas company, *each* of whose customers has the demand curve shown in Figure 9.2. The company charges a high price, P_0 , if the consumer purchases fewer than X units of gas per month. For an amount beyond X units per month, the company charges a medium price, P_1 . For purchases beyond Y , the company charges an even lower price, P_2 . Consequently the company's total revenues from each consumer are equal to the shaded area in Figure 9.2 because

FIGURE 9.2

Second-Degree Price Discrimination

The company charges a different price (P_0 , P_1 , or P_2) depending on how much the consumer purchases, thus increasing its total revenue and profit.



1. Of course, this assumes for simplicity that each consumer purchases Z units and that each price considered exceeds the firm's marginal cost. Also, other simplifying assumptions (which need not concern us here) are made in this and the next paragraph.

the consumer purchases X units at a price of P_0 , $(Y - X)$ units at a price of P_1 , and $(Z - Y)$ units at a price of P_2 .

The manager, by charging different prices for various amounts of the commodity, increases revenues and profit. After all, if she charged only a single price and wanted to sell Z units, she would charge a price of P_2 . Thus the firm's total revenue would equal the rectangle OP_2EZ , which is considerably less than the shaded area in Figure 9.2. By charging different prices, managers increase profit relative to a single-price strategy. Unlike first-degree price discrimination, managers leave a consumer surplus of $A + B + C$. Because second-degree (and third-degree) discrimination occurs at the group level and not at that of the individual, consumers retain some surplus.

Third-Degree Price Discrimination

We now consider the most common form of price discrimination: third-degree price discrimination. Three conditions must hold true for this pricing strategy to succeed. Demand must be heterogeneous, managers must be able to identify and segregate the different segments, and markets must be successfully sealed. As we previously discussed, individuals within a market have different preferences toward a product. The differences in their price elasticity of demand may be due to differences among classes in income levels, tastes, or the availability of substitutes. For example, the price elasticity of demand for the boxes of candy discussed earlier may be lower (in absolute value) for the rich than for the poor.

Think of it like this: Managers would prefer to identify the preferences of individuals (first-degree price discrimination). For various possible reasons, they can't (or don't want to because it is too expensive). So they choose the next best alternative, which is to identify individuals with similar traits and group them together. Managers then appeal to the group.

Students are a good example of third-degree discrimination. Students have relatively limited income, so they tend to have high price elasticities of demand—they are price sensitive. Thus many times they are sold a good or a service at a lower price. They get a discount, and all they must do is to show student identification.

Buyers of the product must fall into classes with considerable differences in price elasticity of demand for the product. Managers must then identify and segregate these classes at moderate cost. Also, buyers must be unable to transfer the product easily from one class to another; otherwise people could make money by buying the product from the low-price classes and selling it to the high-price classes, making it difficult to maintain the price differentials among classes. We call these latter two conditions the ability to *segment* and *seal* the market.²

If managers want to use a third-degree strategy, they must decide how much output to allocate to each class of buyer, and at what price. Suppose there are only two classes of buyers. Managers have already chosen total output, so they need to allocate output across the two markets. Managers will maximize profit by allocat-

2. Segmenting and sealing can have another meaning. In 2000, customers of Amazon.com discovered (via an Internet chat room) that they had been charged significantly different prices by Amazon for the same DVD. When they expressed their displeasure and made the price differences public, Amazon announced that it no longer would engage in such pricing. If the customers had not discovered the price differences, they would have been satisfied (as revealed by their purchase of the DVD), and Amazon could have continued selling the same product at different prices. See David Streitfeld, "On the Web, Price Tags Blur," *Washington Post*, September 27, 2000.

STRATEGY SESSION: That Darling Little Mouse Is Really a Price Discriminator

It's been a long, cold winter in Green Bay. Let's reward the kids with a spring break vacation at Disneyland and Disney California Adventure Park. They'll love Mickey, Minnie, Donald, and Snow White, and it'll be warm. Family A of four (two adults, two kids ages five and seven) packs up and heads for Anaheim.

It's another day in paradise. Should we go to the ocean, take a ride to the mountains, or go to Disneyland and Disney California Adventure Park? So many choices for the Los Angeles family, and so many times those same choices are available. Family B of four (two adults, two kids ages five and seven) hops in the family car and heads for Anaheim.

Both families buy the One-Day Park Hopper (which lets them visit both parks). The bill for the Griswolds from Green Bay is \$312. The bill for the family from Los Angeles is \$292. It costs Disney the same to serve the Los Angeles family as the Green Bay family. So why the \$20 price difference? And how can Disney

tell the two families apart? It's the Wisconsin driver's license that family A is carrying and the driver's license showing a Southern California address that family B is carrying. That's how the market is sealed. How is it segmented?

The Griswolds came all the way from Green Bay to see Mickey, and Dad's not going to disappoint those children for a mere \$20. On the other hand, family B could have gone to the ocean or the mountains or could see Mickey tomorrow. Simply put, the Griswolds had a much less elastic demand than family B, who had plenty of substitutes.

That mouse is a clever third-degree price discriminator!

Sources: http://disneyland.disney.go.com/disneyland/en_US/reserve/ticketListing?year=2007 for Southern California prices and http://disneyland.disney.go.com/disneyland/en_US/reserve/ticketListing?name=TicketListin for other prices.

ing the total output so that the marginal revenue in one class is equal to the marginal revenue in the other. For example, if the marginal revenue in the first class is \$25 and that in the second class is \$10, the allocation is not optimal. Managers can increase profit by allocating one less unit of output to the second class and one more unit to the first class. In fact, managers want to allocate so the marginal revenue of both classes is equal. When this is true, the ratio of the price in the first class to that in the second class equals

$$\frac{P_1}{P_2} = \left[\frac{1 - \left(\frac{1}{|\eta_2|} \right)}{1 - \left(\frac{1}{|\eta_1|} \right)} \right] \quad (9.1)$$

where η_1 is the price elasticity of demand in the first class and η_2 is that in the second class.³ We can now see why it does not pay to discriminate if the two price elasticities are equal: $|\eta_1| = |\eta_2|$ implies that $P_1 = P_2$. Moreover, segments with a lower (absolute values) price elasticity are charged a higher price.

3. Recall from equation (2.15) that marginal revenue equals $P[1 + (1/\eta)]$, where P is price and η is the price elasticity of demand. Therefore, if marginal revenue is the same in the two classes, $P_1[1 - (1/|\eta_1|)] = P_2[1 - (1/|\eta_2|)]$. Hence $P_1/P_2 = [1 - (1/|\eta_2|)]/[1 - (1/|\eta_1|)]$.

Turning to the more realistic case in which managers choose total output, it is obvious they must look at costs as well as demand in the two classes. The manager will then optimize profit when the marginal cost of the entire output is equal to the common value of the marginal revenue in the two classes. The firm's profit (π) is

$$\pi = TR_1 + TR_2 - TC$$

where TR_1 is the total revenue from class 1, TR_2 is the total revenue from class 2, and TC is the total cost. The total cost is a function of the total amount of the good (Q) produced and sold, and it is allocated Q_1 to class 1 and Q_2 to class 2.

The monopolist has two output choices, so profit is maximized when $\Delta\pi/\Delta Q_1 = 0$ and $\Delta\pi/\Delta Q_2 = 0$. Note that $\Delta\pi/\Delta Q_1 = (\Delta TR_1/\Delta Q_1) - (\Delta TC/\Delta Q_1)$ and $(\Delta TR_2/\Delta Q_1) = 0$ because revenues in class 2 are independent of sales in class 1. Likewise, $\Delta\pi/\Delta Q_2 = (\Delta TR_2/\Delta Q_2) - (\Delta TC/\Delta Q_2)$ and $(\Delta TR_1/\Delta Q_2) = 0$ because revenues in class 1 are independent of sales in class 2. These two relationships are rewritten as

$$\begin{aligned}\Delta\pi/\Delta Q_1 &= MR_1 - MC = 0 \\ \Delta\pi/\Delta Q_2 &= MR_2 - MC = 0\end{aligned}\tag{9.2}$$

Note that both $\Delta TC/\Delta Q_1$ and $\Delta TC/\Delta Q_2$ equal MC (and not MC_1 and MC_2) because the plant manager knows only that producing another unit incurs additional costs. It is the marketing or sales department's job to decide whether the good is destined for class 1 or class 2 demanders.

QUANT OPTION

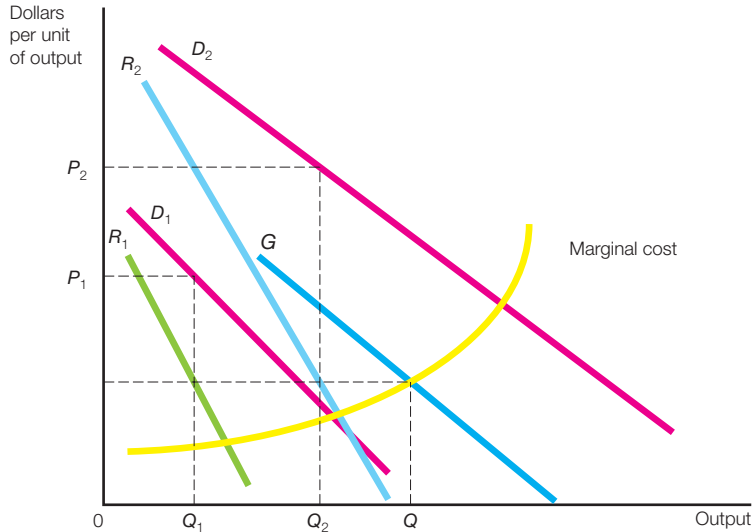
Profits are maximized when $\partial\pi/\partial Q_1 = 0$ and when $\partial\pi/\partial Q_2 = 0$. Partial derivatives are used when just one of the outputs is changing but not the other. Note that $\partial\pi/\partial Q_1 = \{\partial TR_1/\partial Q_1\} - \{\partial TC/\partial Q_1\}$ and $\{\partial TR_2/\partial Q_1\} = 0$ because revenues in class 2 are independent of sales in class 1. Likewise, $\partial\pi/\partial Q_2 = \{\partial TR_2/\partial Q_2\} - \{\partial TC/\partial Q_2\}$ and $\{\partial TR_1/\partial Q_2\} = 0$ because revenues in class 1 are independent of sales in class 2.

The equations (9.2) state that to maximize profit, managers must choose output so $MR_1 = MC$ and $MR_2 = MC$, implying that $MR_1 = MR_2 = MC$. Had there been n classes of demanders, the profit-maximizing rule would be $MR_1 = MR_2 = \dots = MR_n = MC$. To see this in the two-class case, consider Figure 9.3, which shows D_1 , the demand curve in class 1; D_2 , the demand curve in class 2; R_1 , the marginal revenue curve in class 1; R_2 , the marginal revenue curve in class 2; and the firm's marginal cost curve. The curve representing the horizontal summation of the two marginal revenue curves is G . This curve shows, for each

FIGURE 9.3

Third-Degree Price Discrimination

To maximize profit, the firm produces a total output of Q units and sets a price of P_1 in the class 1 market and P_2 in the class 2 market.



level of marginal revenue, the total output needed if marginal revenue in each class is to be maintained at this level. The optimal output is shown by the point where the G curve intersects the marginal cost curve because marginal cost must be equal to the common value of the marginal revenue in each class. If this were not true, profit could be increased by expanding output (if marginal cost were less than marginal revenue) or contracting output (if marginal cost were greater than the marginal revenue). Therefore, managers produce an output of Q units and sell Q_1 units in the class 1 market and Q_2 units in the class 2 market. The price is P_1 in the class 1 market and P_2 in the class 2 market. This results in a higher profit than if the firm quoted the same price in both markets.

Managerial Use of Third-Degree Price Discrimination

Perhaps the most frequently cited example of third-degree price discrimination is the case of airline tickets. The airlines often charge a lower fare for essentially the same ticket if it is purchased in advance as opposed to the day of the flight or based on the day of the week for which the ticket is valid. However, there is a penalty paid on these low-price tickets if the trip is canceled or changed. As an example of different prices for the same service, in 2011 the price of a round-trip coach ticket from New York to San Francisco ranged from about \$580 to \$674 for flights leaving and returning on similar dates.

STRATEGY SESSION: Mickey Mouse Pricing at Amusement Parks

Going to an amusement park used to mean paying the admission charge (and perhaps an additional fee inside the park to ride the hot new attraction); but that was it. The prices were transparent. But now you need a scorecard to know how to play and what prices are available. Attendance fell after 9/11. Despite that, most parks raised their base gate prices.

But in general, that gate price is the highest price to get into a park. There are web-based sales with 17–27% discounts for those who purchase on the web and print their tickets. Others offer further discounts for the observant buyer. For instance, the gate price at Knott's Berry Farm in California is \$43, \$35 online, and \$28.95 on its all-you-can-eat barbecue days. King's Dominion north of Richmond, Virginia, has a \$43.99 gate price, a \$34.99 online price, and a \$29.99 four-day advance purchase price. Five rival parks in California (Disney and Knott's included) teamed together to sell City Pass, which allowed admission to all five parks at a substantial discount. Multiday and annual passes are also generally available only online.

Some of the discounts are cost-based. Online tickets save a park money by not having more sales personnel at the gates and eliminate queues at entrances that annoy customers.

The parks are also exploring other types of tickets. For instance, some people do not want to use a park's attractions but would rather watch others use them—such as grandparents who wish to see their grandchildren enjoy the attractions but who would rather skip riding the “Rebel Yell” roller coaster. The full price may deter onlookers from coming to the park, but a reduced price can mean additional profit because they don't contribute much to costs, are likely to consume high markup concessions and souvenirs, and might preclude the whole family from attending if they didn't attend.

More diverse pricing structures can be based on both the cost and demand sides of the equation.

Source: Eleena De Lisser, “A New Twist in Theme Park Pricing,” *The Wall Street Journal*, June 24, 2004, at online.wsj.com/article/SB108802974024445871.html.

One reason for these price differences is that the price elasticity of demand for business travel is much less elastic than that for vacation travel. Business travelers must meet with clients, suppliers, and associates at particular times, often as soon as possible. Regardless of the price of an airline ticket (so long as it remains within reasonable bounds), many of these trips are well worth making. On the other hand, vacation travelers often plan their trips well in advance, are relatively flexible with regard to the timing of their trips, and are sensitive to moderate differences in ticket price. From the discussion in the previous section, it seems likely that airline managers, to maximize profit, would like to set higher prices for business travelers than for vacation travelers. And this is the effect of the price differences just cited because business travelers are much less likely than vacation travelers to buy their tickets ahead of time and they desire the flexibility of being able to change their flight schedules.

STRATEGY SESSION: Yield Management and Airline Performance

A recent survey has shown that yield management is a major factor in airline profitability. The managers of American Airlines used yield models to generate over \$1 billion in savings over a three-year period. Yield management models are a nice example of how mathematical models apparently capture the complexity of our social structure, using only a few variables. They also show how the intellectual effort of managers generates profit for the firm.

Yield management models are complex pricing mechanisms. They are dynamic in the sense that prices respond to customer behavior. At any time several classes of seats are priced at different levels. Prices at each level depend on a real-time demand forecasting model that analyzes market behavior and then optimizes pricing behavior. The firm prices as if it were practicing third-degree price discrimination.

The models can handle the complexity of reality only by looking at a simplified version of it—as if life is abridged. Yield management models focus on a few key variables and ignore everything else. Most focus on overbooking, discount allocation, and traffic management. Managers build models that look at the revenue potential of a complex menu of price and itinerary pairs.

For example, consider overbooking. Airlines must overbook because some customers never claim their reservations. If they did not overbook, then some aircraft that should fly full based on demand (customer behavior) would fly with unused capacity. So managers build models that balance the trade-offs between the increased revenue of more passengers and the costs of having passengers take the next flight (ideally).

Modeling the situation is not easy. Clearly reputational costs with customers are involved if overbooking becomes too common. There are also real economic costs. Passengers not permitted to board

the plane because of overbooking must be compensated. Many are given vouchers for discounts on future flights, and some must be fed or given hotel rooms.

Managers build the models to maximize expected net revenue. The optimal overbooking rule is to overbook until the expected marginal revenue from one more passenger on a flight is equal to the marginal cost of an additional overbooking. The actual point chosen reflects concerns for customer satisfaction, so it is constrained a bit.

Other variables are modeled using similar decision rules. For example, in discount allocation models, the objective is to balance the expected marginal revenue of a specific fare request against the expected marginal net revenue of all other fares. To see how sophisticated the models are, consider the inclusion of “sell-up” probabilities. These probabilities are used in discounting models to predict which customers will buy a higher-priced ticket if they are not offered a low price.

The models require sophisticated hardware and software to operate efficiently. Computerized reservation systems like SABRE play an integral part in yield management. These systems interface with the market; their ability to capture and analyze data allows the models to constantly update pricing levels. They also control seat inventory.

Given the competitive nature of the airline market, airlines have not been able to keep all the surplus generated by yield management programs. The programs generate some benefits for airline customers, mainly in the form of lower ticket prices and more efficient use of equipment.

Sources: “Yield Management—A Growth Key Driver”; “Airline Ties Profitability Yield to Management,” *The Travel Tightwad*, May 28, 2002.

At the same time, it is also worth noting that because the airlines can reduce their costs if demand is predictable (as a result of better scheduling of equipment and personnel), they may enjoy savings if travelers buy their tickets in advance. Also, if a ticket is not refundable, it clearly benefits the airline more than a ticket that is refundable, even though the penalty involved in changing it may be relatively small.

In recent years entrepreneurs have stepped in with a business model that mitigates some of the airlines' ability to practice third-degree price discrimination. Internet firms such as Expedia scour airline databases continuously looking for cheap fares. Consumers use Expedia and its competitors, such as CheapTicket or Travelocity, to view such fares. These additional information sources can lead to lower fares than if the customer dealt with the air carrier alone. This is not always the case, however. In addition, customers must take the time to search the sites. Because airlines change fares continuously, if the sites are not updated frequently, the customer may not get the cheapest fare. A recent search of two such sites visited within seconds of each other revealed a \$500 difference between the cheapest fares from Philadelphia to Hyderabad, India. Because the airlines release the sale of these seats to the sites, they are still falling into a managerial pricing plan (that plan, however, would be different if the sites were not present). In addition, some sites, like Priceline, follow more of a first-degree strategy. Consumers are asked to name their target price (say X). If the cost of the seat to Priceline was price Y (less than X), then managers have created a surplus of $X - Y$.

USING COUPONS AND REBATES FOR PRICE DISCRIMINATION

One way managers can implement a price discrimination strategy is with coupons and rebates. Basically these devices reduce the price of products. But why don't managers simply reduce prices? Primarily because coupons are used to price discriminate. Not all consumers use coupons. Of 332 billion coupons distributed in the United States in 2010, only 3.3 billion were redeemed. Of consumers, 78.3% reported coupon use in 2010.⁴ This demand segment is more price sensitive and on the more elastic part of the demand curve. Hence managers use coupons and rebates to price discriminate because other consumers (on the less elastic part of the demand curve) are willing to pay more—that is, to buy the good without a coupon.

By estimating the elasticity of demand, managers can figure out how coupons should be priced. Suppose managers at the Barnegat Light Fish Company sell their product, a special blend of crab cake, in a market where managers think two types of consumers exist: a more affluent group (R) with an estimated price elasticity for Barnegat Light crab cakes of -2 ($|\eta_R| = 2$) and a less affluent group (S) with an estimated price elasticity for Barnegat Light crab cakes of -5 ($|\eta_S| = 5$). Managers at the fish company choose a posted price (P) but then issue a coupon for $\$X$

4. Santella and Associates, *Coupon Trends Report*, 2011, www.santella.com/Trends.htm.

off in the newspaper local to the consumer types. Every buyer pays the nominal price of P per unit for Barnegat Light crab cakes on the grocers' sales receipt; but at the bottom of the sales receipt, an $\$X$ credit appears for those who tender a coupon. Thus although all buyers pay the same price P , in reality buyers without coupons pay P while coupon tenderers pay $P - X$. What should the values of P and X be? As we saw, to maximize profit, the marginal revenue in each market should be equal and they, in turn, should equal Barnegat Light's marginal cost (MC). Therefore,

$$P[1 - (1/|\eta_R|)] = (P - X)[1 - (1/|\eta_S|)] = MC$$

Suppose Barnegat Light's marginal cost is a constant \$2

$$MR_R = P[1 - (1/2)] = P/2 = 2 = MC \text{ or } P = \$4$$

and

$$MR_S = (4 - X)[1 - (1/5)] = (4 - X)(0.8) = 2 = MC$$

or

$$3.2 - 0.8X = 2 \text{ or } X = \$1.5$$

Managers should price crab cakes at \$4 per unit and offer a \$1.50 off coupon. The more affluent buyers pay \$4/unit for the crab cakes, and the less affluent ones clip the coupon and pay \$2.50/unit for the *same* crab cakes. Those who are more price elastic (the less affluent in this case) use coupons; the less price elastic people (the affluent in this case) do not. So by issuing coupons (or rebates) managers can price discriminate (and increase their profit).

PEAK LOAD PRICING

The demand for goods or services may shift with the time of day, week, or year. For example, the demand for highway and transit services is greatest during the morning and evening rush hours, lower during midday, and lower still overnight. Roads to resorts are likely to see greater demand on the weekend than during the week. And Miami Beach hotels have greater demand in February when it is cold in the northern United States than in the summer when it's warm almost everywhere in the United States.

Because these temporal differences in demand are coupled with a plant capacity that does not change over the demand cycle, managers facing these demand conditions should charge different prices in the peak (high ones = P_p) and in the trough (low ones = P_t). The rule to follow is that marginal revenue equals marginal cost. However, the marginal revenue curves differ because the service demand curves change between the peak and trough. The marginal cost is usually high in the peak because the supplier is operating at or

PROBLEM SOLVED: Third-Degree Price Discrimination

To illustrate how price discrimination is used, suppose a drug manufacturer sells a major drug in Europe and the United States. Because of legal restrictions, the drug cannot be bought in one country and sold in another. The demand curve for the drug in Europe is

$$P_E = 10 - Q_E \quad (9.3)$$

where P_E is the price (in dollars per pound) in Europe and Q_E is the amount (in millions of pounds) sold there. The demand curve for the drug in the United States is

$$P_U = 20 - 1.5Q_U \quad (9.4)$$

where P_U is the price (in dollars per pound) in the United States and Q_U is the amount (in millions of pounds) sold there. The total cost (in millions of dollars) of producing the drug for sale worldwide is

$$TC = 4 + 2(Q_E + Q_U) \quad (9.5)$$

The firm's total profit (π) from both Europe and the United States is

$$\begin{aligned} \pi &= P_E Q_E + P_U Q_U - TC \\ &= (10 - Q_E)Q_E + [20 - 1.5Q_U]Q_U - [4 + 2(Q_E + Q_U)] \\ &= 10Q_E - Q_E^2 + 20Q_U - 1.5Q_U^2 - 4 - 2Q_E - 2Q_U \\ &= -4 + 8Q_E - Q_E^2 + 18Q_U - 1.5Q_U^2 \end{aligned} \quad (9.6)$$

To maximize profit with respect to Q_E and Q_U , we must set $\Delta\pi/\Delta Q_E = 0$ and $\Delta\pi/\Delta Q_U = 0$. Hence $\Delta\pi/\Delta Q_E = 8 - 2Q_E = 0$ and $\Delta\pi/\Delta Q_U = 18 - 3Q_U = 0$.

Solving these equations for Q_E and Q_U , we find that managers should sell 4 million pounds of the

drug in Europe and 6 million pounds in the United States.

To find the optimal prices in Europe and the United States, we substitute 4 for Q_E and 6 for Q_U in equations (9.3) and (9.4); the result is that managers set a European price of \$6 per pound and a U.S. price of \$11 per pound. Substituting these values of P_E and P_U , as well as the foregoing values of Q_E and Q_U , into equation (9.6), we find that the firm's profit equals

$$\pi = -4 + 8(4) - 4^2 + 18(6) - 1.5(6^2) = 66$$

or \$66 million.

Note that if we use the graphical technique shown in the previous section, we will obtain identical results. Whether the graphical technique or the mathematical technique is used, the answer is the same.

How much additional profit do managers generate? If price discrimination were not possible (perhaps because the submarkets could not be segmented and sealed), P_E would equal P_U . Letting this common price be P , it follows from equation (9.3) that $Q_E = 10 - P$, and from equation (9.4) that $Q_U = (1/1.5)(20 - P) = (40/3) - (2/3)P$. Therefore, the firm's total amount sold in Europe and the United States combined is

$$\begin{aligned} Q &= Q_E + Q_U = (30/3) - (3/3)P + (40/3) - (2/3)P \\ &= (70/3) - (5/3)P \end{aligned}$$

which implies that^a

$$P = 14 - 0.6Q \quad (9.7)$$

QUANT OPTION

To be elegantly fashionable, setting $\partial\pi/\partial Q_E = 0$ and $\partial\pi/\partial Q_U = 0$ will maximize profit. Hence $\partial\pi/\partial Q_E = 8 - 2Q_E = 0$ and $\partial\pi/\partial Q_U = 18 - 3Q_U = 0$.

for $P \leq \$10$ or for $Q \geq 20/3$. (For $P \geq \$10$ or $Q \leq 20/3$, $P = 20 - 1.5Q$ because only the United States purchases the drug if the price exceeds \$10.) Hence managers generate profit of only

$$\begin{aligned}\pi &= PQ - TC \\ &= (14 - 0.6Q)Q - (4 + 2Q) \\ &= 14Q - 0.6Q^2 - 4 - 2Q \\ &= -4 + 12Q - 0.6Q^2\end{aligned}\quad (9.8)$$

because $Q = Q_E + Q_U$.

To maximize profit, the manager selects Q so that $\Delta\pi/\Delta Q = 0$. Therefore,

$$\Delta\pi/\Delta Q = 12 - 1.2Q$$

Solving for Q , we find if managers do not engage in price discrimination, they choose output of 10 million pounds of the drug (which is the same as the output produced when they discriminated).^b Substituting 10 for Q in equations (9.7) and (9.8), it follows that

$$\begin{aligned}P &= 14 - 0.6(10) = \$8 \\ \pi &= -4 + 12(10) - 0.6(10^2) = \$56\end{aligned}$$

Therefore, if managers do not engage in price discrimination, profit is \$56 million rather than the \$66 million they earn by using price discrimination.

Because 10 million pounds are produced under both pricing schemes, the cost of production is the same in both cases: $4 + 2(10) = \$24$. So the total revenues are \$80 ($= 64 + 16$) when there is no discrimination and \$90 ($= 66 + 24$) when there is discrimination. With no discrimination, the average revenue per unit

is just the price ($80/10 = \$8$); but with discrimination, the average revenue per unit is \$9 ($90/10$). The profit-enhancing property of third-degree discrimination is that it raises the average revenue above the price on the demand curve for a given quantity.

If segmenting and sealing the market are possible but costly, the preceding example tells us that managers should be willing to pay up to the difference in the profit of the two pricing schemes (but no more) to segment and seal—that is, up to \$10 million.

Finally, note that at a price of \$6, 4 million pounds of the drug are sold in Europe (from $Q_E = 10 - 6$) and that at a price of \$11, 6 million pounds are sold in the United States (from $Q_U = (40/3) - (2/3)(11)$). Also, note that $\Delta Q_E/\Delta P_E = -1$ and $\Delta Q_U/\Delta P_U = -2/3$. Recall from Chapter 2 that elasticity is $|\eta| = (P/Q)(\Delta Q/\Delta P)$. So $|\eta_E| = (6/4)(|-1|) = |-1.5|$ and $|\eta_U| = (11/6)(|-2/3|) = |-1.22|$. Therefore, the price is raised (from \$8 to \$11) for the less elastic demander and lowered (from \$8 to \$6) for the more elastic demander—just as we would expect from equation (9.1).

^a This means that $(5/3)P = (70/3) - (3/3)Q$ or $5P = 70 - 3Q$ or $P = 14 - 0.6Q$.

^b If the demand curves are curvilinear, the output of the third-degree discriminator and the single-price monopolist are not necessarily the same, and it's possible that prices fall in all markets.

QUANT OPTION

Touché!

Setting $d\pi/dQ = 0$ will maximize profit. Thus, $d\pi/dQ = 12 - 1.2Q$.

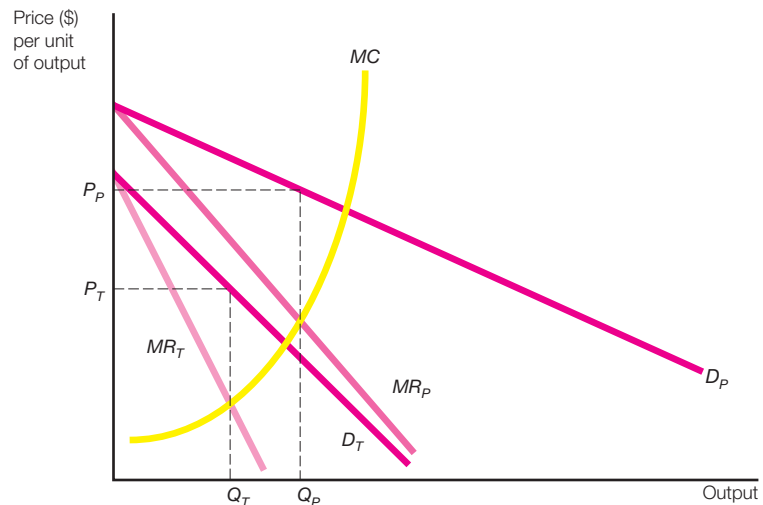
near capacity, and it is usually low in the trough because much excess capacity exists. Note that this is not the same as third-degree price discrimination. Both the third-degree price discrimination and peak trough situations have separate marginal revenues for each demand class; but in third-degree price discrimination, the demand classes share the *same* supplier capacity at the *same* time. Therefore, marginal cost in third-degree price discrimination is a function of $Q_1 + Q_2$; that is, the two demands are interdependent in how they influence marginal cost. But in the intertemporal demand case, the demanders use the *same* capacity at *different* times. Therefore, there are separate levels of marginal cost for Q_1 and Q_2 ; that is, the demands are independent in their influence on marginal cost. The optimal solution for third-degree price discrimination is $MR_1(Q_1) = MR_2(Q_2) = MC(Q_1 + Q_2)$, whereas the optimal solution for peak–trough pricing is $MR_1(Q_1) = MC_1(Q_1)$ and $MR_2(Q_2) = MC_2(Q_2)$. The parentheses indicate “a function of.” These conditions are shown in Figure 9.4.

The Strategy Session box discusses peaks and troughs in electricity demand. Consider the situation on roadways. The Texas Transportation Institute reports that American drivers in the largest 101 urban areas on average lost 40 hours in 2010 to road congestion.⁵ This indicates severe auto congestion in some areas. The worst is in Washington, DC, where the typical driver could save 74 hours

FIGURE 9.4

Determination of Peak and Trough Prices

The optimal peak price (P_P) is determined by where the peak marginal revenue equals the firm’s marginal cost; the optimal trough price (P_T) is determined by where the trough marginal revenue equals the firm’s marginal cost.



5. David Schrank and Tim Lomax, *The 2011 Urban Mobility Report* (College Station: Texas Transportation Institute, Texas A & M University, September 2011).

STRATEGY SESSION: The Future Is Now: The Futures Market for Super Bowl Tickets

Your team is going to the Super Bowl. You'd like to go. You could have bought tickets way in advance of the game, but you didn't know your team was going until they won their league championship game (and you're only interested in being there if your team is playing). So you go online and see if you can buy tickets on eBay, or you call a ticket broker, or you go to the Super Bowl city on game day and seek a ticket scalper outside the stadium. You could have bought the tickets before you knew who was in the game and then sold them on eBay if your team did not make the game.

But a new market has grown up for you. Commodities have had futures markets for years. Now there's a futures market for Super Bowl tickets. In such markets both buyers and sellers can lock in a price and reduce uncertainty. If you wait until your team makes it, you don't know what the ticket price will be.

Yoonew.com, TheTicketReserve.com, and SuperbowlOption.com sold options on the 2006 Super Bowl. A client pays \$X for the option of getting a ticket for the game. If his or her team makes the Super Bowl, he or she gets a ticket. If not, the option is worthless. The futures contracts are for a specific team, so the price will vary from team to team based on the likelihood of that team making the Super Bowl. Prices change as more information becomes available. For instance, if a team clinches their division crown, that guarantees them a slot in the playoffs (and increases their chances of being in the Super Bowl). This raises the price of their option. If the team wins the first round of their divisional playoffs, this puts them a step closer to the Super Bowl and hence increases the price of their option. If they win the second-round playoff game, they are in the Super Bowl. Even then the value of the option will change based on the collective demand of fans to see the game (because the

supply is fixed—the stadium has a certain capacity). Data suggests that gambling odds from SportingbetUSA explained 96% of the variability in one options market's prices.

Alan Krueger, an economist at Princeton, gives the following example of how the market works. Suppose there's a 10% chance a fan's team will reach the Super Bowl, and a futures contract costs \$250, whereas a ticket when the Super Bowl participants are known is \$2,500. Note that the expected value of a \$2,500 ticket is \$250 ($0.1 \times \$2,500$). Suppose a risk-loving fan would pay \$2,500 for a ticket to see his team play and a risk-averse fan would pay \$250 for a futures contract. But there's more. A risk-averse fan will be willing to pay more for a futures contract. As with an insurance policy, ticket futures sell at a premium over their expected value because they help risk-averse fans hedge against uncertainty.

What premium? The fan could guarantee a ticket to the Super Bowl by buying a futures contract on every team in a conference. This is a sure thing. If fans were risk-neutral, the sure-thing price would equal the price the tickets were expected to be at game time (say \$2,500). The excess of the sure-thing price over \$2,500 is a measure of the market valuation of insuring against risk. For the 2006 Super Bowl, the premium for a sure-thing ticket ranged from 35 to 60% during the season (which is not far from the risk premium in some lines of regular insurance). However, this premium fell substantially as the playoffs advanced. Krueger hypothesizes that this occurred because fans overestimated their teams' chances of getting in the Super Bowl earlier in the season.

Source: Alan B. Krueger, "Wait Till Next Year, but Lock In the Ticket Price Now," *New York Times*, February 2, 2006, at www.nytimes.com/2006/02/02/business/02scene.html.

per year if he or she could drive at free flow rates on the roadway—that is, at the posted speed limit. One reason why such levels of congestion exist is that roadways, in general, are not peak–trough priced in the United States. Singapore has used peak–trough pricing in its central city since the 1970s. In 2003 London instituted a £5 (now £8, and the pricing area has expanded geographically) price for driving in central London. Initial reports are that driving has decreased by 20%. (Are you surprised by the direction of this change? Having studied managerial economics, we hope you are not.) State Route 91 in Orange County, California, has priced recently constructed lanes (where the price varies in real time to keep the lanes operating at free flow level) and kept the existing lanes free (where rush hour traffic moves at 10–25 mph). Many transit systems run 10 times as much equipment during peak rush hours compared to off-peak periods. However, many transport systems charge a flat fee to use the system, independent of the time of use. In fact, many systems actually reward peak use by selling weekly or monthly passes at a discount compared to purchasing single rides for each commute. Therefore, some peak riders actually pay less per ride compared with off-peak riders—just the opposite of what we stated was optimal.⁶ Some systems, such as the Metro system in Washington, DC, practice peak–trough pricing.

Another version of intertemporal pricing exists. Some consumers have to read the best-selling book on the *New York Times* list as soon as it reaches that position (or perhaps before it reaches that position if they are truly trendsetters). Others must see the latest Harry Potter or Daniel Day-Lewis movie the first weekend it opens (or soon thereafter). Such individuals can discuss the book or movie at the next cocktail party or around the water cooler and be judged “worthy” by their peers. These people have a high demand to be “with it” and hence pay a high price for hardcover books and first-run movies.

Others have an interest in such books or movies but not at the prices that the trendsetters will pay. After about a year, the paperback version of the best seller appears at 20–40% of the price of the hardcover book. And after about six months, the DVD of the movie is available for purchase for less than two admissions to the first-run movie theater (and you can see it again and again, pause while you do something else, and rewind to see a favorite scene).

So book and movie suppliers realize that there are leaders and followers in the markets for their services and have figured out how to cater to both with high prices for those who cannot wait and low prices for those who can.

Two-part tariff When managers set prices so that consumers pay an entry fee and then a use fee for each unit of the product they consume.

6. Such a pricing policy may be related to the fact that transit’s competition (the highway system) has a zero price during peaks and troughs.

TWO-PART TARIFFS

Often managers will implement a first-degree price discrimination strategy through a **two-part tariff**. Managers set prices so that consumers pay an *entry fee* and then a *use fee* for each unit of the product they consume. Two-part tariffs are

STRATEGY SESSION: Why Do Your Laundry at 3 A.M.?

Gertrude Stein once wrote that “a rose is a rose is a rose.” In most states a kilowatt-hour is a kilowatt-hour—but not in Florida, Pennsylvania, Washington, and Wisconsin. These states allow electric utilities to practice time-of-day pricing to residential customers who opt to be charged in this way. Otherwise, consumers can stick with the traditional plan, where they pay the same flat rate per kilowatt-hour all the time for power. In the new pricing schemes, prices of a kilowatt-hour can change every several hours. (A Pennsylvania utility, Allegheny Power, has experimented with rates that change every hour.) Not surprisingly, in these states it costs more to consume a kilowatt-hour in the peak periods, when demand is greatest, and less during the trough periods, when demand is the lowest. Many states have allowed time-of-day pricing for commercial and industrial customers for quite some time.

In Florida, Gulf Power of Pensacola charges \$0.042 cents per kilowatt-hour at night, on weekends, and on holidays. Demand is less during those periods because the 9-to-5 workday crowd is not at work. Gulf charges \$0.10 cents per kilowatt-hour on weekday afternoons when residential and commercial power demand peaks because of air conditioning use. A third rate is a “critical rate” of \$0.309 cents when supplies of kilowatt-hours go

extremely short (less than 1% of the time). These rates compare to the alternative residential plan of a flat \$0.063 cents per kilowatt-hour regardless of the time. A customer utilizing the plan estimates that he shaved \$600 off his annual power bill by shifting a third of his power consumption to the off-peak periods. In Washington, Puget Sound Energy has estimated that running the same dishwasher in the off-peak time saves a user 25% off the peak rate. With about one-third of customers participating in the off-peak plan, Puget Energy Inc. (the biggest residential time-of-use provider) estimates that peak demand has been cut by 5%. This saves the energy company big money. If it cannot handle peak loads, it enters the power grid market to buy the required power at spot market rates (which are usually expensive) or brings its least efficient (most expensive) capacity on line. By restricting the quantity demanded in the peak via pricing, it need not resort to these expensive alternatives. In addition, by bolstering trough demand, it utilizes its capital plants better.

Source: R. Gavin, “Cut Your Electric Bill: Do Laundry at 3 A.M.,” *The Wall Street Journal*, August 22, 2002.

common in the business world. Membership fees for golf clubs are the entry fee, and a greens fee for playing a round of golf is the use fee. Wireless phone users are asked to pay an initial fee and then are charged monthly fees in exchange for access to the network. Some are even charged a use fee for each message unit. Other examples include razors and blades, health clubs, and computer printers. An innovative (and lucrative) use of a two-part tariff is the personal seat license (PSL) for sports stadiums. While the stadium is being constructed, fans are asked to pay a PSL. This is a fixed fee (usually thousands of dollars) that gives the fan the right to purchase tickets to attend a game. It is also an important source of revenue for

STRATEGY SESSION: A Change from Markup Pricing to Sophisticated Pricing

Parker Hannifin produces over 800,000 individual parts. Many (about a third) are virtually one of a kind with limited or no competition. How should these items be priced? Until 2001 the answer was cost plus (the cost of making the part plus a 35% markup). Sophisticated computer models costed out an item and then added on 35% (approximately, with some discretion given to sales where competition and hence price comparisons were easier for clients).

This pricing mechanism created several results counter to the best profit interests of the firm. First, if Parker became more efficient and reduced its costs, it automatically reduced its price. This might make sense if competitors were lowering their prices or if lowering Parker's prices would yield more profitable business; but if there was limited competition, this was a missed profit opportunity. Second, if Parker improved its product at the same cost and thus added value to the customer, Parker maintained its price and got nothing for its value-enhanced product. It's estimated by Thomas Nagle of the Monitor Group that as many as 60% of U.S. manufacturers use cost-plus pricing.

When Donald Washkewicz became CEO of Parker in 2001, he decided to practice strategic pricing (defined as basing prices on determining what a customer was willing to pay as opposed to what it costs to make a product). Since Parker implemented strategic pricing, operating income increased by \$200 million, net income increased by \$543 million, return on invested capital increased to 21% (from 7%), and Parker's share prices rose by 88% (all in the 2002–2006 time frame).

Washkewicz views the change of corporate culture as being like pulling teeth. His vice president of corporate strategic pricing defines it as messing with the company's DNA. Washkewicz had to terminate

some executives who couldn't get with the new program. Now each of the company's 115 divisions has at least one pricing guru to implement its strategic pricing. One guru describes the previous cost-plus pricing policy as one where no one asked, Why not a 45% markup?

Parker has divided its 800,000 products into five categories. A core product is highly competitive with many external reference prices. Prices in this category fell modestly (in some cases by 3%) but also increased (in some cases by 5%). Two classes (*B* and *C*) of partially differentiated (from the market) products exist. In *B* the differentiation adds value to the customer; prices here increased by 0–5%. In *C* the products were niche in nature with no close competitors; here prices increased by 0–9%. The differentiated product systems were tailored to improve customer profitability and productivity, and prices in this category increased by 0–25%. The last category was custom designed, and only Parker could do it. Here prices increased by over 25%.

Although some customers balked at the price increases, virtually all stuck with Parker, especially because Parker promoted the value-added properties of its products to customers. Adopting strategic pricing also has impacted how Parker thinks about new product development. One dimension now considered is the ability of each product to yield a pricing premium.

With the financial gains shown by Parker, one wonders what the other 60% of companies that practice cost-plus pricing are thinking.

Source: Timothy Aepfel, "Seeking Perfect Prices, CEO Tears Up the Rules," *The Wall Street Journal*, March 27, 2007, p. 1.

construction costs. In effect, managers use PSLs to generate revenue from an asset that hasn't yet been built.

One decision facing managers is to set the appropriate fixed upfront fee and variable usage fee to maximize profit. The upfront fee is designed to extract consumer surplus, so managers use it for first-degree price discrimination. One example is a country club (which actually practices a three-part tariff). Before you can play a round of golf, you must be approved for membership. With that comes a one-time initiation fee, then yearly dues. Both payments are made before you can play and are independent of the number of rounds you play. In this sense the initiation fee and the dues are like the fixed costs managers face. But should you wish to play a round of golf, having been selected as a member and being a member in good standing, you must also pay a greens fee (that is, a use fee for the service the club provides). The greens fee is analogous to the variable costs managers face. Eating clubs, tennis clubs, health clubs, and amusement parks all practice similar pricing policies. In some amusement parks, one fee (the entry fee) gets you inside the park, where the fee to go on the rides (the use fee) is zero for many rides; but some rides (the newest or the most popular) often require additional fees for rides taken.

Managers in other markets also use two-part tariffs. Walmart's Sam's Club is one example. After paying a membership fee, members are admitted to a Sam's Club store, where they pay individually for every item purchased. As another example, it is estimated that Costco earns over 50% of its profit from its entry fee. This bodes well for Costco customers because it means individual items are being priced close to marginal cost.

Managers at Internet service providers also use this pricing strategy. For a fixed monthly fee (the entry fee), customers get access to the Internet. Then they are charged for each time unit they are online (the use fee). In many cases the use fee is zero for the first X minutes; but after the X minutes, a per-minute fee is assessed.

We start with a simple example to demonstrate this pricing principle. Suppose all demanders for a service are perfect clones; *each* demander has the same demand curve. That is, they all have identical preferences. We assume the demand curve is linear of the form $P = a - bQ$, where P is the price per unit and Q is the number of units demanded at price P . In addition, we assume managers face a constant marginal cost of production.

The profit-maximizing optimal two-part tariff requires pricing the use fee at marginal cost and the entry fee equal to the resulting consumer surplus. So managers must choose their use fee before pricing the entry fee. Consider the situation in Figure 9.5. The use fee (P^*) equals MC . At P^* , the demander consumes Q^* units. The resulting consumer surplus A^* is the optimal entry fee.⁷ The use fee covers the manager's variable cost of serving the consumer (because $MC = AVC$ and $(AVC)Q^* = VC$) when marginal costs are constant, and the variable-cost profit

7. Technically, $A^* - \varepsilon$ will break the indifference of the consumer to joining or not joining. Hereafter we'll call it A^* , recognizing that the nonambiguous entry fee is $A^* - \varepsilon$.

STRATEGY SESSION: Making Them Pay Twice: Personal Seat Licenses for Sports Teams

Charlotte, North Carolina, is a city on the move. It is the banking capital of the southeastern United States. One way that many upcoming cities “get on the map” is to obtain a professional sports franchise. Charlotte acquired a National Football League franchise for the Carolina Panthers. But it needed a stadium. How should it finance such a large capital expense?

Enter Max Muhlemann. Charlotte was excited about its new team, and fans were supportive. Muhlemann suggested that the Panthers sell personal seat licenses for the new stadium. The concept was that a fan would have to purchase a personal seat license to be able to purchase a ticket to see the Panthers play football. The personal seat license was an entry fee. The use fee was the price of a game ticket.

Demand for the licenses was strong; they sold for prices that reflected the desirability of the seats. The average price of a personal seat license was \$2,400. The Panthers sold 62,500 personal seat licenses. That’s \$149 million received by the team *before* a game was ever played in the stadium. That’s \$149 million in consumer surplus that fans were willing to spend just for the *right* to purchase tickets to see the games. After obtaining the personal seat license, the holder would buy a game ticket at a price no different than other football teams charge. The big difference was that other teams were not collecting the consumer surplus as the Panthers were.

The use of personal seat licenses to finance new stadiums and stadium improvements is growing. The New York Jets and the New York Giants are building a \$1.6 billion new stadium to be shared by both teams in New Jersey. Although this has not yet been announced, virtually all commentators and fans expect both teams to use personal seat licenses to help finance the stadium. If they sell them, it is expected that all 9,200 club seats (the category below luxury suites) would be sold as personal seat licenses, along with a number of other seats. Twelve NFL teams have used

seat licenses since the mid-1990s and have raised almost \$900 million in the process. The Philadelphia Eagles sold 29,000 such licenses at prices ranging from \$1,800 to \$3,700 and raised \$70 million toward the cost of their new field. Dallas has sold one for its yet-to-be-built stadium at \$150,000. Other Dallas luxury seats are going for \$16,000, \$35,000, \$50,000, and \$100,000. The Cowboys’ senior vice president for sales and marketing has said that “it was an internal feeling that that was what the market could bear”—which is what price discrimination is all about. Half of the Dallas licenses sold in a four-month period. It is estimated that the Cowboys could raise \$300 million. The Chicago Bears sold 45% of their seats with the highest price being \$10,000. The remaining 55% of the seats carried no licenses.

Although some fans complain about the licenses because of an active resale market, other fans (in cities where attendance and demand for tickets are high) view the licenses as an appreciating asset. Some Chicago fans (of the 55% who did not have to purchase licenses) approached the Bears *and requested that their seats be licensed*. The Bears did so. The licenses are resold in many ways. One is via Seasonticketrights.com. Its founder reports that the average gain on the resale of a Bears license is about \$8,300 (which shows that the Bears underpriced their licenses). To demonstrate that fans have a higher reservation price than they are currently charged, one Giants fan who currently pays \$80 per game ticket says, “I’m going to buy my tickets whether there are licenses or not. Do I want to pay? No.” But he says he will.

Source: F. Klein, “Growing Plague: Buying the Right to Buy a Ticket,” *The Wall Street Journal*, September 26, 1996; and Richard Sandomir, “Jets and Giants Fans May Pay for the Right to Pay for Tickets,” *New York Times*, March 22, 2008, at [www.nytimes.com/2008/03/22/sports/football/22seat.html?scp=1&sq=jets+and+giants+fans+may+pay+for+the+right+to+pay+for+tickets"&st=nyt](http://www.nytimes.com/2008/03/22/sports/football/22seat.html?scp=1&sq=jets+and+giants+fans+may+pay+for+the+right+to+pay+for+tickets).

STRATEGY SESSION: Costco and the Two-Part Tariff

"People laughed at the idea of charging someone to shop at your warehouse, but our membership fees are north of \$1 billion per year," states Joel Benoit, a senior vice president at Costco. There are more than 24 million members in the United States and Canada. Current yearly membership fees are \$50 per person, per family, or per business and \$100 for an executive membership (which entitles the customer to other services). Note that 24 million members at \$50 per member is \$1.2 billion, and because some are executive members, that \$1.2 billion is a conservative estimate.

Steve Hoch, a professor of marketing at the Wharton School, states that most of Costco's profit is

from the annual membership fee (the entry fee). But profit also occurs from the markup of items' costs. There is a scale advantage here too: The larger the Costco membership, the larger the item quantities Costco buys. That greater buying power gives Costco a lower cost of obtaining items and the ability to offer items at lower prices.

Source: Julie Bick, "24 Rolls of Toilet Paper, a Tub of Salsa, and a Plasma TV," *New York Times*, January 28, 2007, at [www.nytimes.com/2007/01/28/business/yourmoney/28costco.html?scp=1&sq="24+rolls+of+toilet+paper"&st=nyt](http://www.nytimes.com/2007/01/28/business/yourmoney/28costco.html?scp=1&sq=).

of the firm for serving this consumer is $A^* + P^*Q^* - (AVC)Q^* = A^*$ (because $AVC = P^*$). Multiplying A^* by the number of clones and subtracting the firm's fixed cost gives managers their profit.

Intuitively, a two-part tariff lets managers act as first-degree price discriminators. Managers capture the entire consumer surplus through the entry fee and convert it into producer surplus (variable-cost profit). Note that managers produce until price equals marginal cost.

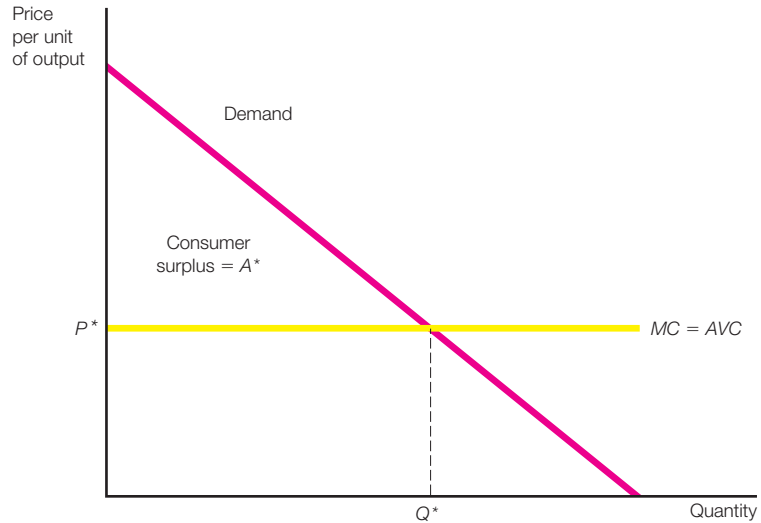
A two-part tariff is simpler for managers to implement than first-degree price discrimination because they need not charge individuals different prices for each unit of the good consumed. This pricing strategy also gives managers two other advantages. First, the entry fee is collected at the beginning of the demand period. First-degree price discrimination collects as the consumer consumes the product or service. Money now is worth more than the same amount of money later. Second, though most managers understand there are some units that are highly valued (the first round of golf after a long Minnesota winter) and some units whose value to consumers is low (a round of golf in a downpour), they are not certain of individuals' reservation prices. With the two-part tariff, all a manager must know is that over the demand period, consumers will show variance in their reservation prices. By using a two-part tariff strategy, managers need not worry about this variance. They have already collected their surplus with the entry fee.

Managers also use two-part tariffs to get customers to reveal their preferences. Because it is costly to induce customers to reveal their demand functions, many

FIGURE 9.5

Optimal Two-Part Tariff When All Demanders Are the Same

The optimal two-part tariff when all demanders are clones is a use fee equal to marginal cost ($P^* = MC$) and an entry fee equal to the consumer surplus resulting from such a use fee (A^*).



managers offer a pricing menu to customers. Consider the wireless phone industry. Managers offer customers different plans that vary in monthly charges and use fees. Customers choose the plans they believe are optimal for them; hence they reveal their preferences.

A Two-Part Tariff with a Rising Marginal Cost

What if managers face marginal costs that are upward-sloping rather than constant? The optimal rule for managers remains the same: Charge a use fee equal to marginal cost and an entry fee equal to the resulting consumer surplus. The only difference, relative to constant marginal cost, is that managers realize additional profit from the use fee (see area X^*) as well as their entry fee, as shown in Figure 9.6.

Charging a use fee of P^* results in selling Q^* to the consumer. This yields revenues of $P^*Q^* = X^* + Y^*$ from the use fee. The variable cost of selling Q^* units to the customer is the area under the marginal cost curve (Y^*). Therefore, the revenue from the use fee more than covers the variable-costs of serving the customer, and managers earn a variable-cost profit from serving the customer of X^* from the use fee. The entry fee is the consumer surplus that results from charging the use fee of P^* (that is, A^*). Hence the variable-cost profit of serving this customer is $A^* + X^*$.

STRATEGY SESSION: Verizon Local Calling Plans

Telephone service is a classic example of a two-part tariff. A subscriber pays the phone company a monthly fee for the privilege of receiving a dial tone. This fee must be paid regardless of whether there are zero, tens, or hundreds of incoming or outgoing calls per month. It is an entry fee for having the service available (although one could avoid this fee by using pay phones, albeit at a higher use fee for outgoing calls).

Managers of phone companies are becoming more adept at using sophisticated pricing strategies. Consider Verizon's local calling plans in New Jersey. Because Verizon managers do not know, with certainty, the demand curve of individual consumers, they let the consumers reveal their preference function by offering a menu of pricing plans. Although most of the plans are primarily two-part tariff pricing, managers combine this pricing strategy with bundling and price discrimination.

In 2008 Verizon offered the following local calling plans in New Jersey:

- *Low-use message rate service*—\$5.20/month: This plan has a monthly allowance of 20 message units per month with every unit over 20 costing \$0.10/unit. A message unit is a local call of five minutes.
- *Moderate message rate service*—\$7.40/month: This plan has a monthly allowance of 75 message units per month with every unit over 75 costing \$0.065/unit.

- *Flat-rate service*—\$8.95/month: This plan gives the consumer unlimited message units of local outgoing calls during the month.
- *Verizon local package*—\$25.99/month: This plan gives the consumer unlimited message units of local outgoing calls during the month and a choice of up to three calling features.
- *Verizon local service package extra*—\$29.99/month: This plan gives the consumer unlimited message units of local outgoing calls during the month and a choice of four or more calling features.
- *Verizon regional package unlimited*—\$38.00/month: This plan gives the consumer unlimited message units of local outgoing calls, unlimited message units of regional outgoing calls, and a choice of up to three calling features.
- *Verizon regional package*—\$42.95/month: This plan gives the consumer unlimited message units of local outgoing calls, unlimited message units of regional outgoing calls, and five calling features including home voice mail.

Special features include unlimited directory assistance and calling features. Features include caller ID, three-way calling, and call waiting.

Source: www.verizon.com, accessed on March 24, 2008.

A Two-Part Tariff with Different Demand Curves

In most markets consumers do not all have identical demand functions. What is the optimal two-part tariff when there are multiple types of demanders in the market? Consider the case with relatively strong and weak demanders. The strong demander is willing to purchase more units than the weak at any given price. Managers should consider at least two two-part tariff pricing options. If

PROBLEM SOLVED: Two-Part Tariff Pricing

Let's demonstrate the use of a two-part tariff pricing strategy. Managers at C-Pal Industries face 100 identical individuals, each with a demand curve of $P = 10 - Q$. C-Pal has a constant marginal cost of \$4 per unit produced and a fixed cost of \$500. C-Pal's situation is depicted in the figure.

Managers at C-Pal charge a use fee of \$4 ($= MC$) for each good a consumer purchases. Consumers purchase six goods apiece; the demand can be rewritten as $Q = 10 - P = 10 - 4 = 6$. C-Pal's total revenue from the use fee from one customer is $P^*Q^* = (\$4)(6) = \24 , and C-Pal's variable cost for serving one customer is $(AVC)Q^* = (\$4)(6) = \24 . The consumer surplus when six goods are demanded at a price of \$4 for a customer is $0.5(10 - 4)6 = \$18$, and managers charge this as an entry fee. The total revenue from one customer is $\$24 + \$18 = \$42$, and the variable cost of serving that customer is \$24, yielding C-Pal a variable-cost profit of serving a customer of $\$42 - \$24 = \$18$; this is the consumer surplus captured from the consumer and converted into producer surplus. Because there are 100 clones, managers earn a total

variable-cost profit of $100(\$18) = \$1,800$. C-Pal's profit is the variable-cost profit minus the fixed costs: $\$1,300 = \$1,800 - \$500$.

One point of confusion in using a two-part tariff is what happens when a demander conceives of a two-part tariff as a one-part tariff. Consider one of C-Pal's customers. He is paying (on average) \$7 for each item he consumes—that is, \$4 from the use fee and \$3 ($= \$18/6$) from the entry fee. But if C-Pal had merely put a flat charge on each item sold of \$7, the customer would purchase only three units ($Q = 10 - P = 10 - 7 = 3$). The individual demand curve derived in Chapter 3 shows the amount the consumer pays for *each* unit. Indeed, if C-Pal's customers face a price of \$7, they will purchase only three units. But that is not the deal they have been offered. The *only* way they can buy the good is to pay an entry fee of \$18 for the privilege of purchasing each unit at a price of \$4. They choose to purchase six because their benefits equal their costs of doing so. That is why the two-part tariff is so clever. It extracts all the consumer surplus (which a single price does not).

A Two-Part Tariff Example: C-Pal Industries

C-Pal's optimal two-part tariff entails charging a use fee of \$4 ($= MC$) for each item consumed and an entry fee of \$18.

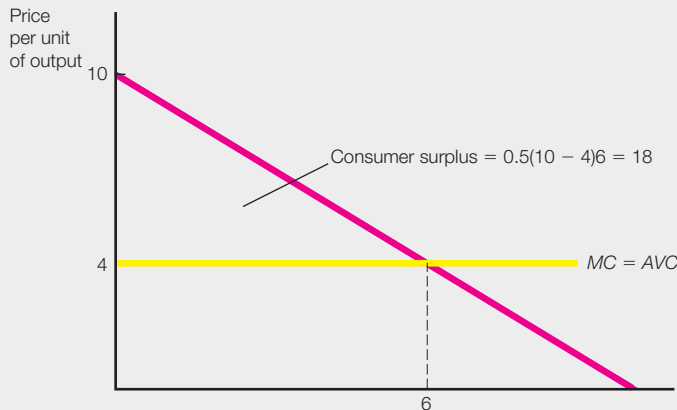
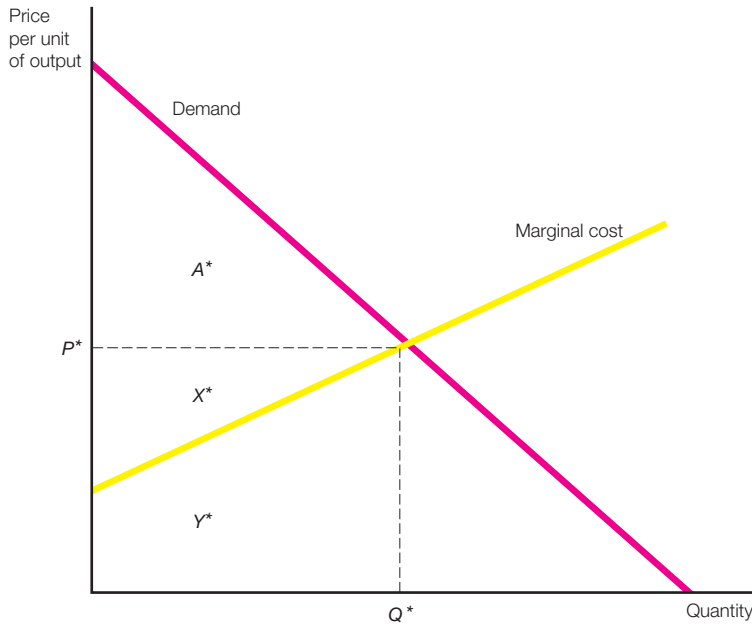


FIGURE 9.6

Optimal Two-Part Tariff When Marginal Cost Is Rising

The optimal two-part tariff is to charge a use fee P^* equal to marginal cost and an entry fee equal to the resulting consumer surplus (A^*). The firm's variable-cost profit is now $A^* + X^*$ because the firm's use revenues now exceed its variable cost (Y^*).



the strong demander is willing to buy significantly more units at any price, then it is more profitable to charge a use fee equal to marginal cost and an entry fee equal to the resulting consumer surplus of the strong demander. This strategy excludes the weak demander from the market. The weak demander's consumer surplus is smaller than that of the strong demander and so the weak demander is not willing to pay the entry fee. From her point of view, the marginal cost (entry fee) is greater than the marginal benefit (consumer surplus). It is not unusual for managers to use pricing policies that exclude demanders from markets. In single-price scenario, no consumer whose reservation price is below the market price participates in that market.

The other pricing policy is used when the strong demand is not that much stronger than the weak demand. In these markets, managers should set the use fee at or above marginal cost and set the entry fee equal to the resulting consumer surplus of the weak demander. In doing so, managers cannot use first-degree price

STRATEGY SESSION: Academic Institutions Practice What They Preach

While the professional teams have opted for the two-part tariff pricing structure called “personal seat licenses” (in which a fan purchases a seat license—the entry fee—for the right to purchase game tickets—the use fee), some universities, such as the University of California at Berkeley and the University of Kansas, have opted for a variant. With personal seat licenses, the price of game tickets can vary from year to year. The university version (called Equity Seats Rights, or ESR) has an entry fee and a use fee, but the ESR fee guarantees the ESR holder that the use fee (ticket price) won’t go up for a certain period of time (10 years in the case of KU and 30 years in the case of UC Berkeley). KU’s rights could cost as much as \$105,000, while UC Berkeley’s rights could cost as much as \$225,000. The rights can be paid in a lump sum or in installments that are likened to a mortgage—that is, like a \$105,000 mortgage with yearly installments of principal and interest due.

What’s the advantage to the buyer over a PSA? The first advantage is the guaranteed use fee. The

second is that 70% of the purchase price can be deducted from federal taxes because it’s treated as a contribution to a nonprofit institution. KU projects that about 20% of the expected revenues from its ESR will be spent on academic programs, with the remaining 80% going to athletic programs. UC Berkeley had sold 1,800 of 3,000 available seats by March 2010 and had raised \$150 million. KU has been less successful.

Will the pro teams go in the same direction? They haven’t so far. The Chicago Cubs baseball team nixed the concept. The Sacramento Kings basketball team was considering the concept as was the Tottenham (English Premier League) soccer team.

Sources: David Sweet, “Cal’s Unique Seat-pricing Concept,” *sportsbizonMSNBC.com*, November 19, 2008, at www.msnbc.com/id/27788392/ns/business_sports_biz/t/new_seating_concept_could_net_cal_million/#.Ts8P13HWNsQ; and Associated Press, “Colleges’ Mortgage’ Top Seats,” March 30, 2010, at http://cjonline.com/sports/2010-03-30/colleges_mortgage_top_seats#.Ts8QGxHWNsQ.

discrimination against the strong demander and this demander will realize some consumer surplus. The situation is depicted in Figure 9.7

If managers want to exclude the weak demander, they should set the use fee equal to marginal cost ($= AVC$) and the entry fee equal to the relevant consumer surplus of the strong demander. The revenue from the use fee equals the variable cost incurred serving the strong demander. The variable-cost profit is the entry fee (areas A^* through F). If managers want to include the weak demander, they must choose the use fee P^* , which maximizes the area $2A^* + 2C + D + E$ or $2A^* + 2C + 2D$, whichever is larger. $2A^* + 2C + D + E$ occurs if $P^* > MC$ and $2A^* + 2C + 2D$ occurs if $P^* = MC$. Once P^* is chosen, it determines the consumer surplus (either A^* or $A^* + C + D$).

STRATEGY SESSION: Scientific Pricing—Even for Great Art?

A number of firms have been practicing scientific pricing for some time. Airlines have milked their databases to learn the booking trends on routes each hour until takeoff and how those booking rates respond to price changes at various times. They also factor in forward-looking information such as conventions and major sporting events. Long's Drug Stores and D'Agostino Supermarkets have developed prices for specific products at specific stores (and prices for item X may differ from store to store even if the stores are close). But pricing salsa based on ethnic characteristics has got to be different than pricing Picassos, doesn't it?

Apparently not, according to David Galenson, an economist (and art lover and collector) from the University of Chicago. Galenson has developed a model to explain the value of great art, and it works fairly well. After collecting price data on works of great artists, Galenson found the following pattern. Young great artists (Gauguin, Picasso, Van Gogh) seemed to have epiphanies (something came to them and they put it down on canvas quickly). At the other end of the spectrum were the old great artists (like Cezanne), whose great paintings were modifications and evolutions of previous work. Galenson claims the same poles exist for novelists, too.

While economists understand how to model human behavior (or so we think) and use data to test

the model, behaviorists feel that human behavior is too complex to be captured in a regression equation. This type of debate also rages in sports. Many professional sports teams have computerized every play of every game and look for patterns in coaches' play calling in certain situations to predict the behavior of a coach in a current situation. In addition, they document players' performances in every play. As an example, they know whether baseball player X tends to hit better against left-handed or right-handed pitchers (and which pitchers) and how he performs in pressure situations. Old-time scouts say you need a feel for the game that a computer can't give you. Some doctors also abhor "evidence-based medicine," in which data point to the diagnosis and remedy; they prefer a doctor's clinical judgment.

In the meantime, don't overpay for your next work of great art. That Picasso, painted by the old man, isn't worth much relative to the young Picasso's work; and that painting Cezanne did as a young man isn't worth as much as the work of the mature Cezanne.

Source: David Leonhardt, "The Art of Pricing Great Art," *New York Times*, November 15, 2006, at www.nytimes.com/2006/11/15/business/15leonhardt.html.

Because both demanders are willing to pay A^* in the case of $P^* > MC$, managers realize $2A^*$ in revenues. The revenues from the use fee more than cover the variable cost of serving the consumers. At the use fee of $P^*(> MC)$, the weak demander wants Q_w units of the good and the strong demander wants Q_s units of the good. Area C represents the variable-cost profit managers realize from the use fee revenues from the weak demander, and area $C + D + E$ represents the

PROBLEM SOLVED: A Two-Part Tariff with Different Demands

The Will and Dylan Company has a strong demander (with a demand curve of $P_S = 8 - Q_S$) and a weak demander (with a demand curve of $P_W = 6 - Q_W$). Managers face a constant marginal cost of production of \$2. They want to consider several two-part tariff pricing options in order to increase firm value. They first consider charging a use fee of \$2 (the firm's marginal cost) and an entry fee equal to the resulting consumer surplus of the strong demander. We can rewrite the strong demand curve as $Q_S = 8 - P_S$. If the use fee is \$2, the strong demander will purchase six units. The resulting consumer surplus is $0.5[8 - 2]6 = \$18$. The managers choose this as the entry fee. Under this strategy, managers earn a variable-cost profit of \$18. Managers next consider charging a use fee of \$2 and an entry fee equal to the consumer surplus of the weak demander. Because we can rewrite the weak demand as $Q_W = 6 - P_W$, if the use fee is \$2, the weak demander will purchase four units. The resulting consumer surplus for the weak demander equals $0.5[6 - 2]4 = \$8$. If this is charged as the entry fee, both demand types will pay it, and the firm's variable-cost profit will be \$16.

Finally, managers consider charging a use fee greater than marginal cost and an entry fee equal to the resulting consumer surplus of the weak demander. How should managers choose the optimal use fee [P^*]? If managers charge a use fee of P^* , the strong demander will purchase $Q_S = 8 - P^*$ units and the weak demander will purchase $Q_W = 6 - P^*$ units. Because $P^* > MC = AVC$, managers will realize a variable-cost profit (of $P^* - 2$) from every unit they sell (and they sell $8 - P^* + 6 - P^* = 14 - 2P^*$ units). The variable-cost profit from the use fee is $(P^* - 2)(14 - 2P^*) = -2P^{*2} + 18P^* - 28$. With $6 - P^*$ units sold to weak demanders, their resulting consumer surplus is $0.5[6 - P^*][6 - P^*] = 18 - 6P^* + 0.5P^{*2}$. This is the entry fee; and because both demanders will pay it, the variable-cost profit from the entry fee is $36 - 12P^* + P^{*2}$. The total variable cost profit then is

$$\begin{aligned} VC\pi &= -2P^{*2} + 18P^* - 28 + 36 - 12P^* + P^{*2} \\ &= -P^{*2} + 6P^* + 8 \end{aligned} \quad (9.9)$$

The variable-cost profit is maximized when $\Delta VC\pi / \Delta P^* = 0$. Therefore,

variable-cost profit realized from the use fee revenues from the strong demander. Therefore, the total variable-cost profit is $2A^* + 2C + D + E$ from serving both demander types if $P^* > MC$.

If the use fee P^* is set equal to marginal cost, the resulting consumer surplus of the weak demander is $A^* + C + D$. Both demanders will pay it. There is no profit from the use fee because it equals marginal cost (= average variable cost). The profit managers earn is thus $2A^* + 2C + 2D$. We must then compare $2A^* + 2C + 2D$ with $2A^* + 2C + D + E$ to see which is largest. That determines (if we serve both demanders) whether the use fee is greater than or equal to marginal cost. Once we determine the best profit from serving both demanders, we must compare it with the profit from serving just the strong

$$\Delta VC\pi / \Delta P^* = 0 = -2P^* + 6$$

or

$$P^* = \$3$$

Substituting $P^* = \$3$ in equation (9.9) gives

$$VC\pi = -(3^2) + 6(3) + 8 = \$17$$

Managers compare the resulting variable-cost profit from each strategy, and choose to serve the strong demander only. This gives them the highest variable-cost profit of \$18.^{ab}

The managers consider one last pricing option. Suppose they combine the concept of price discrimination and the two-part tariff. They calculated the consumer surplus of the weak demander when the use fee is a marginal cost (= \$2) of \$8, and that of the strong demander when the use fee is marginal cost which is \$18. Therefore, they propose to set the use fee equal to marginal cost, charge the weak demander an entry fee of \$8, and charge the strong demander an entry fee of \$18. This yields a variable-cost profit of \$26. Price discrimination takes place not on the use fees but on the entry fees.

Think about this last pricing policy in the real world. Clubs have full members, associate members, junior members, and the like. Each has a different initiation and dues structure. Usually there is some restriction on use (perhaps not all members can play golf on Wednesday afternoons, when the doctors play). But can you see the motivation behind these different classes of membership?

^a If the strong demand had been $P_s = 7 - Q_s$, serving only the strong demander would yield a variable-cost profit of \$12.5, whereas serving both demand types with a use fee equal to marginal cost would yield a variable-cost profit of \$16. If a use fee greater than marginal cost is chosen (optimal fee \$2.5) and the resulting consumer surplus of the weak demander is the entry fee (\$6.125), the variable-cost profit is \$16.25, which is the best of the three options considered.

^b We showed the case of $P^* = MC = 2$ explicitly. We did not need to do this. If $P^* = MC = 2$ was the profit-maximizing use fee, it would be the solution to the maximization of the variable-cost profit equation (9.9).

QUANT OPTION

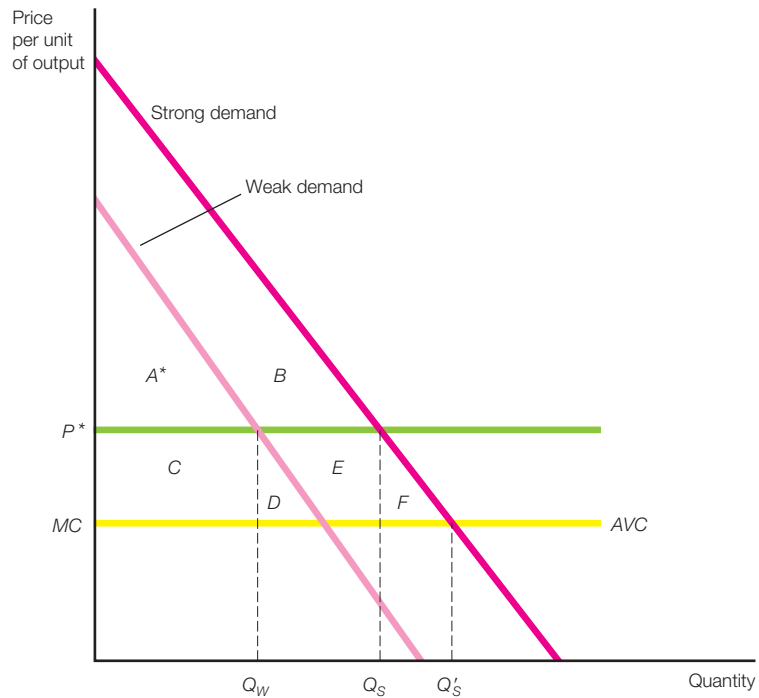
Setting $dVC\pi / dP^* = 0$ will maximize profit. Thus, $dVC\pi / dP^* = -2P^* + 6 = 0$.

demander. When we solve this problem with the calculus (see Problem Solved), the two calculations for the weak demander collapse to one and the calculus shows whether $P^* > MC$ or $P^* = MC$ maximizes profit if the manager serves both demanders.

FIGURE 9.7

Optimal Two-Part Tariff with Two Demand Types

The use fee should be set equal to marginal cost and the entry fee equal to the resulting consumer surplus of the strong demander (areas A^* through F) if areas A^* through F exceed the maximum of $2A^* + 2C + D + E$ and $2A^* + 2C + 2D$. The use fee should be set equal to $P^*(>MC)$ and the entry fee equal to the resulting consumer surplus of the weak demander (A^*) if $2A^* + 2C + D + E$ exceeds areas A^* through F or equal to $P^* = MC$ and the entry fee equal to the resulting consumer surplus of the weak demander ($A^* + C + D$) if the above condition holds and $2A^* + 2C + 2D$ exceeds $2A^* + 2C + D + E$.



Consider the best possible scenario for managers: perfectly price discriminate on the basis of entry fees. Managers will charge all demanders a usage fee equal to marginal cost. Weak demanders are willing to pay an entry fee equal to $A^* + C + D$. Stronger demanders will pay an entry fee of $A^* + C + D + B + E + F$. Managers then realize profits of $2A^* + 2C + 2D + B + E + F$. All consumer surplus is converted to producer surplus. We see managers using this strategy. For example, health clubs may have silver, gold, platinum, and diamond membership fees.

Two-part tariff pricing schemes can get complicated fairly quickly. If demand curves intersect, the analysis becomes more complex than our discussion here. We treat this more difficult case in the chapter appendix. Fortunately, complex cases can be modeled so that a number of demanders with varying demand characteristics are considered. Managers need to understand the potential increase in profit due to two-part tariff pricing. Several available models examine more complex demand in markets.

SUMMARY

1. Managers practice price discrimination either when they sell physically identical products at different prices or when similar products are sold at prices with different ratios to marginal cost. The strategy works best in markets with various classes of buyers who are differentiated in price elasticities of demand; where segments can be identified and segregated with relatively low costs (lower than the added expected revenue); and where markets can be sealed so goods cannot be transferred easily from one class to another. Once managers choose a discriminating strategy, they maximize profit by allocating outputs across markets so the marginal revenues are equal to each other and to the total marginal cost. This is called third-degree price discrimination. Managers use second-degree price discrimination when they can price increments of output at different rates, usually charging higher rates for initial increments of output, then lower rates as consumption increases. First-degree price discrimination entails pricing goods at the reservation price of each consumer. This practice captures all the consumer surplus and converts it to producer surplus or variable-cost profit. It is the strategy of first choice because of this. However, it is difficult to estimate each consumer's reservation price, and this scheme is more costly to implement than the other degrees of price discrimination.

2. Two-part tariffs are a strategy to enable managers to use first-degree price discrimination. This pricing strategy has managers charge the consumer an "entry" fee for the right to pay a "use" fee to actually purchase the product. In the simplest case, where all demanders are the same, the optimal use fee is the marginal cost of the product and the entry fee is the consumer surplus available when that use fee is charged. If consumers have different demand curves, managers may exclude weaker demanders from the market and follow the preceding rule with the stronger demanders. Or managers could include all demanders by pricing the use fee at or above its marginal cost and choosing an entry fee equal to the resulting consumer surplus of the weak demander. Managers who practice price discrimination on the entry fee, while charging all consumers the marginal cost as a use fee, realize the maximum profit.

3. Consumer preferences tend to show temporal variation (by day, week, or season). To account for these variations in temporal behavioral, managers many times charge high prices during the peaks and lower prices during the troughs (as opposed to a single price across the whole temporal cycle). The rule for managers to optimize price is to set the relevant marginal revenue equal to marginal cost.

wwwnorton.com/studyspace 

PROBLEMS

1. Managers at the Ridgeway Corporation produce a medical device that they sell in Japan, Europe, and the United States. Transportation costs are a negligible proportion of the product's total costs. The price elasticity of demand for the product is -4.0 in Japan, -2.0 in the United States, and -1.33 in Europe. Because of legal limitations, this medical device, once sold to a customer in one country, cannot be resold to a buyer in another country.
 - a. The firm's vice president for marketing circulates a memo recommending that the price of the device be \$1,000 in Japan, \$2,000 in the United States, and \$3,000 in Europe. Comment on his recommendations.
 - b. His recommendations are accepted. Sales managers send reports to corporate headquarters saying that the quantity of the devices being sold in the United States is lower than expected. Comment on their reports.
 - c. After considerable argument, the U.S. sales manager agrees to lower the price in the United States to \$1,500. Is this a wise decision? Why or why not?
 - d. Can you be sure that managers are maximizing profit? Why or why not?
2. Ann McCutcheon is hired as a consultant to a firm producing ball bearings. This firm sells in two distinct markets, each of which is completely sealed off from the other. The demand curve for the firm's output in the first market is $P_1 = 160 - 8Q_1$, where P_1 is the price of the product and Q_1 is the amount sold in the first market. The demand curve for the firm's output in the second market is $P_2 = 80 - 2Q_2$, where P_2 is the price of the product and Q_2 is the amount sold in the second market. The firm's marginal cost curve is $5 + Q$, where Q is the firm's entire output (destined for either market). Managers ask Ann McCutcheon to suggest a pricing policy.
 - a. How many units of output should she tell managers to sell in the second market?
 - b. How many units of output should she tell managers to sell in the first market?
 - c. What price should managers charge in each market?

3. The Lone Star Transportation Company hauls coal and manufactured goods. The demand curve for its services by the coal producers is

$$P_C = 495 - 5Q_C$$

where P_C is the price (in dollars) per ton-mile of coal hauled and Q_C is the number of ton-miles of coal hauled (in thousands). The demand curve for its services by the producers of manufactured goods is

$$P_M = 750 - 10Q_M$$

where P_M is the price (in dollars) per ton-mile of manufactured goods hauled, and Q_M is the number of ton-miles of manufactured goods hauled (in thousands). The firm's total cost function is

$$TC = 410 + 8(Q_C + Q_M)$$

where TC is total cost (in thousands of dollars).

- What price should managers charge to haul coal?
 - What price should managers charge to haul manufactured goods?
 - If a regulatory agency were to require managers to charge the same price to haul both coal and manufactured goods, would this reduce the firm's profit? If so, by how much?
4. Electric companies typically have 5–10 different rate schedules for their main customer groups. The average price charged to large industrial users may differ substantially from that charged to residences. Moreover, many consumers pay a price for electricity based on the time of day they use it. For example, the prices charged by Consolidated Edison, a large New York electric utility, and Pacific Gas and Electric, a major California electric utility, are as follows:

Price	
Company and Time of Day of Electricity Use (Cents per Kilowatt-Hour)	
Consolidated Edison	
8 A.M.–10 P.M. (peak hours)	27
10 P.M.–8 A.M. (off-peak hours)	4 ^a
Pacific Gas and Electric	
Summer	
Noon–6 P.M. (peak hours)	28.3
6 P.M.–noon (off-peak hours)	9.2
Winter	
Noon–6 P.M. (peak hours)	11.3
6 P.M.–noon (off-peak hours)	8.0

^aApproximate figure.

Electric utilities use their cheapest generators continuously and start up their more costly ones as demand goes up. Consequently, at 3 A.M., a utility might meet its requirements from a hydroelectric dam that produces electricity for \$0.02 per kilowatt-hour. However, on a hot day in August, when air conditioners are running full blast, demand would be so great that the utility would be forced to use its most costly generators—perhaps an oil-fired plant where electricity costs \$0.07 per kilowatt-hour.

- a. Does price discrimination occur in the market for electricity?
 - b. Why have some state regulatory commissions, including the Public Service Commission of New York, ordered that time-of-day rates be phased in for residential consumers?
 - c. In many areas, both residential and industrial consumers tend to pay a lower price per kilowatt-hour if they use more rather than less electricity. Is this price discrimination? If so, what kind of price discrimination is it?
 - d. Explain why price discrimination is used by managers of electric companies.⁸
5. In the town of Oz, there are two types of tennis players: wizards andimps. Wizards andimps do not socialize, so it would be impossible to start a tennis club that both types would join. Imps have access to credit but a weak demand for tennis as follows.

$$P_I = 30 - Q_I$$

where Q_I refers to the number of games they would play if the price of a game were P_I .

Because of their access to credit, they would be willing to pay an upfront fee to join the club.

Wizards live from paycheck to paycheck and would be willing to pay for each tennis game as they go along. Their demand is

$$P_W = 40 - Q_W$$

where Q_W refers to the number of games they would play if the price of a game were P_W .

There are an equal number of wizards andimps (for simplicity, assume one of each). The marginal cost of one game of tennis is a constant 2.

You can design your tennis facility to attract either wizards orimps (but not both). Which clientele would you like to attract and what would be your profit per “person”?

6. The managers of Roosevelt’s (a local yet upscale bar) are considering charging an admission fee on Thursday nights. They contemplate how to charge. Should they
 - Option 1. Use just a beverage charge per beverage ordered *or*
 - Option 2. Use an admission charge (a fee to enter the establishment) *and* a beverage charge per beverage ordered?

8. For further discussion, see W. Shepherd and C. Wilcox, *Public Policies toward Business* (Homewood, IL: Irwin, 1979); and *New York Times*, June 9, 1990.

There are two types of people who frequent Roosevelt's: Over 21 Students (S) and Over 21 Student Wannabees (W). Each Student has a demand for beverages of

$$P = 8 - Q_S$$

where Q_S is the quantity of beverages demanded if the price of a beverage is P . Each Wannabee has a demand for beverages of

$$P = 8 - 2Q_W$$

where Q_W is the quantity of beverages demanded if the price of a beverage is P .

The marginal cost of serving a beverage is a constant \$2.

For simplicity, assume there is one demander of each type. Roosevelt's must (by law) charge all customers the same admission charge and the same per beverage charge. Beverages do not have to be sold in integer amounts and prices do not have to be in integer amounts.

- a. Under option 1, what is the profit maximizing price per beverage?
 - b. Under option 2, what is the profit maximizing two-part tariff?
 - c. What is Roosevelt's profit under Roosevelt's best choice?
7. The demand for a strong demander for a round of golf is

$$P_S = 6 - Q_S$$

where Q_S is the number of rounds demanded by a strong demander when the price of a round of golf is P_S .

The demand for a weak demander for a round of golf is

$$P_W = 4 - Q_W$$

where Q_W is the number of rounds demanded by a weak demander when the price of a round of golf is P_W .

The cost of providing an additional round of golf to either type of golfer is a constant 2.

There is one golfer of each type.

The club has decided that the best pricing policy is a two-part tariff. However, it's your job to tell the club the optimal entry fee and the optimal use fee to maximize the club's profit. The club cannot price discriminate on either the use or the entry fee. The club's fixed cost is 1.

What are the club's optimal use fee and the optimal entry fee?

8. The university museum has two types of visitors. One type is university employees; and the other type is people nonaffiliated with the university. All university employees have identical annual demands for museum visits, given by

$$P_P = 30 - Q_P \quad (\text{for each university employee})$$

where Q_p is the number of visits demanded if the price is P_p per visit. Non-affiliated people all have identical annual demands for museum visits, but differ from university employees

$$P_N = 100 - Q_N \quad (\text{for each nonaffiliated person})$$

where Q_N is the number of visits demanded if the price is P_N per visit. The museum can identify university employees by their university ID card, while a nonaffiliated person does not possess a university ID.

The university's profit-maximizing museum is contemplating two different pricing policies:

Policy 1

- For university employees: An annual membership fee and an additional price-per-visit. (Only university employees are eligible for this membership plan.)
- For nonaffiliated visitors: A single price-per-visit, with no membership fee. (This price per visit is not necessarily the same as the university employee price per visit.)

Policy 2

- This policy would offer a *different* price-per-visit for each type of visitor, but no membership fees at all.

The museum has a constant marginal cost of \$6 per visit, regardless of the visitor's type. For simplicity, assume that there is one university employee and one nonaffiliated person in the target population.

How much more profit does the best policy yield than the other policy?

EXCEL EXERCISE: PERFECT PRICE DISCRIMINATION

Suppose our monopolist takes their demand curve ($P = 14 - Q$) and charges each demander they wish to serve the buyer's reservation price. Recall that the firm's total cost is $TC = 5 + 2Q + 0.5Q^2$, and the firm's marginal cost is $MC = 2 + Q$. We assume here that the good can be produced and sold on a continuous basis, that is, in non-integer amounts.

The perfectly discriminating firm's demand curve becomes the firm's marginal revenue curve (as shown in the text). Therefore, the firm will produce where marginal revenue (MR) equals MC , that is,

$$MR = 14 - Q = 2 + Q = MC$$

or $2Q = 12$

or $Q = 6$

and the last item sold is to the demander with the reservation price of 8.

The firm receives total revenue equal to the trapezoid running from $P = 14$ at $Q = 0$ to $P = 8$ at $Q = 6$ or

$$TR = 0.5 \cdot (14 + 8) \cdot 6 = 0.5 \cdot 22 \cdot 6 = 66$$

The firm's total cost is

$$TC = 5 + 2Q + 0.5Q^2 = 5 + 2 \cdot 6 + 0.5 \cdot 6 \cdot 6 = 5 + 12 + 18 = 35$$

Thus, the perfectly discriminating firm's profit is

$$\pi = TR - TC = 66 - 35 = 31$$

(substantially better than the 19 shown in the excel exercise on simple monopoly pricing in Chapter 8).

Can we do perfect price discrimination with spreadsheets? Yes we can.

We will assume now that our product can only be produced and consumed in integers.

Call up your spreadsheet. Enter 14 in cell A1, 13 in cell A2, and continue until you enter 0 in cell A15. Then enter 0 in cell B1, 1 in cell B2, and so on, until you enter 14 in cell B15. With columns A and B, you have entered the firm's demand curve.

Then enter 0 in cell C1, =A2 in cell C2, =C2+A3 in cell C3. Then click on the lower right-hand corner of cell C3 with your mouse and drag down until cell C15. This will give you a column of the firm's total revenue at each quantity if you perfectly price discriminate, that is, charge each demander their reservation price.

Then enter =C2-C1 in cell D2, =C3-C2 in cell D3, and so on. You may do this by clicking on the lower right-hand corner of cell D2 and dragging down with your mouse until you reach cell D15. This gives you the firm's marginal revenue curve. Note that the result is exactly the same as column A, that is, the perfectly discriminating monopolist's demand curve and its marginal revenue curve coincide.

Then enter =5 in cell E1, and so on, via click and drag until cell E15. This is a column of the firm's fixed cost.

Then enter =2*B1+0.5*B1^2 in cell F1, and so on via click and drag, until cell F15. This is a column of the firm's variable cost.

Then enter =E1+F1 in cell G1, and so on via click and drag, until cell G15. This is a column of the firm's total cost.

Then enter =G2-G1 in cell H2, and so on via click and drag, until cell H15. This is a column of the firm's discrete marginal costs, the additional cost incurred when producing another unit of product.

From above, we know that the firm's continuous marginal cost is $MC = 2 + Q$. For comparison, we'll enter the continuous marginal cost curve in column I. Enter =2+B1 in cell I1 and so on via click and drag until cell I15. Note

that the two marginal costs differ because of the size of the increments considered (i.e., integer versus continuous), and that the continuous marginal cost exceeds the discrete marginal cost.

Column J is our objective, that is, the firm's profit. Enter $=C1-G1$ in cell J1, and so on via click and drag, until cell J15. Then search column J for the largest number. You may enter $=\text{Max}(J1:J15)$ in cell J16 to have the spreadsheet do this for you. The largest profit will be 28 in cell J7, and as you read leftward in the seventh row, you will notice that this occurs when you produce and sell six units (the same as the above with the continuous demand and cost quantities).

For the continuous case, marginal revenue = 8 and marginal cost = 8 at six units of output. In the discrete case, marginal revenue = 8 and marginal cost = 7.5 at six units of output so the firm wants to produce and sell the sixth unit since the cash flow in (8) exceeds the cash flow out (7.5). The discrete firm does not want to go beyond six units because the marginal cost of the seventh unit = 8.5 exceeds the marginal revenue of the seventh unit = 7. Since cash flow in (7) is exceeded by cash flow out (8.5), the firm will not produce the seventh unit.

Why does the profit differ between the continuous model (31) and the discrete model (28)? This is because of the difference in the changes in output. The continuous demand model is $P = 14 - Q$, whereas the discrete demand function is a step function. Let's look at the continuous demand function from output zero to output one. The total revenue over this demand range is the trapezoid with area of

$$TR = 0.5 \cdot (14 + 13) \cdot 1 = 13.5$$

whereas the total revenue for the discrete demand curve in this demand range is 13 (the additional revenue from going from selling zero units to selling one unit). That's a difference of 0.5. There's an additional 0.5 between outputs one and two, and yet another between outputs two and three. This continues up to output six. There are six of these 0.5 values. Hence, $6 \cdot 0.5 = 3$ which accounts for the 31–28 difference in profits.

EXCEL EXERCISE: THIRD-DEGREE PRICE DISCRIMINATION

Suppose that you know the demand for your product is made up of the demand from strong demanders and the demand from weak demanders. The strong demanders have a demand curve of

$$P_S = 14 - Q_S \text{ or } Q_S = 14 - P_S$$

and the weak demanders have a demand curve of

$$P_W = 10 - Q_W \text{ or } Q_W = 10 - P_W$$

The total demand for your product can be expressed as

$$Q = Q_S + Q_W = 14 - P + 10 - P = 24 - 2P$$

or $2P = 24 - Q$

or $P = 12 - 0.5Q$ for $P \leq 10$ and $Q \geq 4$

The firm's marginal revenue (as shown in the text) has the same intercept as and double the slope of the demand curve, that is,

$$MR = 12 - Q$$

The firm's total cost function is

$$TC = 5 + 2Q$$

The firm's marginal cost is

$$MC = \Delta TC / \Delta Q = 2$$

The firm's average variable cost is

$$AVC = VC / Q = 2Q / 2 = 2$$

To maximize profit, the simple monopoly firm will set $MR = MC$ (as shown in the text), that is,

$$MR = 12 - Q = 2 = MC$$

or $Q = 10$

Substituting $Q = 10$ into the demand curve gives

$$P = 12 - 0.5 \cdot 10 = 12 - 5 = 7$$

Total Revenue = $TR = P \cdot Q = 7 \cdot 10 = 70$

Total Cost = $TC = 5 + 2Q = 5 + 2 \cdot 10 = 5 + 20 = 25$

Profit = $\pi = TR - TC = 70 - 25 = 45$

If the firm followed third-degree price discrimination, they would set $MR_S = MC$ and $MR_W = MC$ (as shown in the text). In other words,

$$MR_S = 14 - 2Q_S = 2 = MC$$

or $2Q_S = 12$

or $Q_S = 6$

Substituting $Q_S = 6$ into the strong demand curve gives

$$P_S = 14 - Q_S = 14 - 6 = 8$$

and $MR_W = 10 - 2Q_W = 2 = MC$

$$\text{or} \quad 2Q_W = 8$$

$$\text{or} \quad Q_W = 4$$

Substituting $Q_W = 4$ into the weak demand curve gives

$$P_W = 10 - Q_W = 10 - 4 = 6$$

Note that $Q = Q_S + Q_W = 6 + 4 = 10$

Total Revenue Strong Market = $TR_S = P_S * Q_S = 8 * 6 = 48$

Total Revenue Weak Market = $TR_W = P_W * Q_W = 6 * 4 = 24$

Total Revenue Total = $TR = TR_S + TR_W = 48 + 24 = 72$

Total Cost = $TC = 5 + 2Q = 5 + 2 * 10 = 5 + 20 = 25$

Profit = $\pi = TR - TC = 72 - 25 = 47$

Thus following third-degree price discrimination increases profit by $47 - 45 = 2$.

Let's suppose we didn't know the calculus derived rules for profit maximization but were given the following estimates of market demand by our demand consultants:

Price	Strong Quantity Demanded	Weak Quantity Demanded
14	0	0
13	1	0
12	2	0
11	3	0
10	4	0
9	5	1
8	6	2
7	7	3
6	8	4
5	9	5
4	10	6
3	11	7
2	12	8
1	13	9
0	14	10

Can we find the simple monopoly price with this information? Yes.

Open your spreadsheet. Enter 14 in cell A1, 13 in cell A2, and so on, until you enter 0 in cell A15. Column A is the price (P) column.

Enter 0 in cell B1, 1 in cell B2, and so on, until you enter 14 in cell B15. Column B is the strong demander's quantity demanded (Q_S).

Enter 0 in cell C1, 0 in cell C2, 0 in cell C3, 0 in cell C4, 0 in cell C5, 1 in cell C6, 2 in cell C7, and so on, until you enter 10 in cell C15. Column C is the weak demander's quantity demanded (Q_w).

Enter $=B1+C1$ in cell D1 and so on. You can do this easily by clicking on the lower right-hand corner of cell D1 and dragging your mouse down until you reach cell D15. Columns A and D give the simple monopolist's demand curve, that is, the total quantity demanded by the strong and weak demanders ($Q = Q_s + Q_w$) at a given price.

Enter $A1 * D1$ in cell E1, and so on, via click and drag until cell E15. Column E gives the simple monopolist's total revenue ($TR = P * Q$).

Enter $=E2-E1$ in cell F2, $=E3-E2$ in cell F3, $=E4-E3$ in cell F4, $=E5-E4$ in cell F5, $=(E6-E5)/2$ in cell F6, and so on via click and drag, until cell F15. Column F is the simple monopolist's marginal revenue, that is, the additional revenue received from selling another unit of product ($MR = \Delta TR / \Delta Q$). The entries in cells F6 to F15 are divided by 2 because quantity demanded changes by 2 when price drops by a dollar below \$10.

Enter 5 in cell G1. Click and drag down to cell G15 to get a column of the firm's fixed cost (FC).

Enter $=2 * D1$ in cell H1, and so on via click and drag, until cell H15. Column H is the firm's variable cost (VC).

Enter $=G1+H1$ in cell I1, and so on via click and drag, until cell I15. Column I is the firm's total cost ($TC = FC + VC$).

Enter $I2-I1$ in cell J2, $=I3-I2$ in cell J3, $=I4-I3$ in cell J4, $=I5-I4$ in cell J5, $=(I6-I5)/2$ in cell J6, and so on via click and drag, until cell J15. Column J is the firm's marginal cost ($MC = \Delta TC / \Delta Q$). The entries in cells J6 to J15 are divided by 2 because quantity changes by 2 when price drops by a dollar below \$10.

Column K is the profit column ($\pi = TR - TC$). Enter $=E1-I1$ in cell K1, and so on via click and drag, until cell K15. Then search column K for the highest number (or enter $=\text{Max}(K1:K15)$ in cell K16 and let the spreadsheet find the highest number). You'll find the maximum profit in the eighth row in cell K8 at 45. This is the same result we attained using the calculus version above.

Now let's do third-degree price discrimination.

Repeat columns A and B starting in cell A17 and B17. Then enter $=A17*B17$ in cell C17, and so on via click and drag, until cell C31. Column C is the total revenue from the strong demand curve.

Enter $=C18-C17$ in cell D18, and so on via click and drag, until cell D31. Column D gives the marginal revenue from the strong demand curve.

Enter $=2*B17$ in cell E17, and so on via click and drag, until cell E31. Column E gives the variable cost of producing product to be sold in the strong market.

Enter $=E18-E17$ in cell F18, and so on via click and drag, until cell F31. Column F gives the marginal cost incurred in producing product to be sold in the strong market.

Enter $=C17-E17$ in cell G17, and so on via click and drag, until cell G31. Column G gives the firm's variable profit, that is, $TR_s - VC_s$, for the strong market. We then search column G for the highest number, which is 36 in cell G23 with output of 6 and a price of 8. This is the same result found with the calculus. We could also find this result by entering $=\text{Max}(G17:G31)$ in cell C32.

Note that the marginal revenue at $Q=6$ is 3 and the marginal cost is 2. From our earlier discussion, we definitely want to sell the sixth unit since the additional cash flowing into the firm (3) exceeds the cash flowing out of the firm (2). The firm would not want to sell a seventh unit in the strong market because the marginal revenue of the seventh unit is 1 and the marginal cost of the seventh unit is 2. Thus, using our rules, six is the optimal number of units for the strong market.

Enter $=-A18/B18$ in cell H18, and so on via click and drag, until cell H31. Column H gives the own price elasticity of demand at each point on the strong demand curve. Note that when the monopolist chose the best single price (7), seven units were sold in the strong market and the demand elasticity was -1 .

Now we do the same for the weak market.

Repeat columns A (the market price) and C (the weak quantity demanded) starting in cells A33 and B33 but beginning with price 10 and quantity 0. Then enter $=A33*B33$ in cell C33, and so on via click and drag, until cell C43. Column C is the total revenue from the weak demand curve.

Enter $=C34-C33$ in cell D34, and so on via click and drag, until cell D43. Column D gives the marginal revenue from the weak demand curve.

Enter $=2*B33$ in cell E33, and so on via click and drag, until cell E43. Column E gives the variable cost of producing product to be sold in the weak market.

Enter $=E34-E33$ in cell F34, and so on via click and drag, until cell F43. Column F gives the marginal cost incurred in producing product to be sold in the weak market.

Enter $=C33-E33$ in cell G33, and so on via click and drag, until cell G43. Column G gives the firm's variable profit, that is, $TR_w - VC_w$, for the weak market. We then search column G for the highest number, which is 16 in cell G37 with output of 4 and a price of 6. This is the same result found with the calculus. We could also find this result by entering $=\text{Max}(G33:G43)$ in cell G44.

Note that these results are similar to those in the strong market. At $Q=4$ marginal revenue is 3 and marginal cost is 2. Whereas at $Q=5$, marginal revenue is 1 and marginal cost is 2. Thus, 4 is the optimal output for managers to sell in the weak market.

Enter $=-A34/B34$ in cell H34, and so on via click and drag, until cell H43. Column H gives the own price elasticity of demand at each point on the weak demand curve. Note that when the monopolist chose the best single price (7), three units were sold in the weak market and the demand elasticity was -2.333 . This demonstrates the result pointed out in the text, that we tend to raise the

price (from 7 to 8) in the least elastic market (strong with elasticity of -1) and lower the price (from 7 to 6) in the most elastic market (weak with elasticity of -2.333).

Adding the two variable profits gives total variable profit ($36 + 16 = 52$). Subtracting the firm's fixed cost of 5 yields the firm's profit of 47, the result of third-degree price discrimination.

APPENDIX: TWO-PART TARIFF WITH INTERSECTING DEMANDS

Managers may find themselves executing a two-part tariff pricing strategy in markets where demand curves intersect. In such a case determining the optimal use price and entry price may entail searching back and forth among the demand curves, as we show in the following example.

There are two types of demanders in our market. Consumer 1 has a demand curve of

$$P_1 = a_1 - b_1Q_1$$

and consumer 2 has a demand curve of

$$P_2 = a_2 - b_2Q_2$$

Marginal cost is constant at c .

Managers can estimate the consumer surplus for any demander i when the use fee is set at marginal cost. Consumer surplus when the use fee is equal to marginal cost is

$$CS_i = (a_i - c)^2 / 2b_i$$

where $i = 1, 2$.

One fact managers can estimate is whether double the consumer surplus for the smaller consumer surplus demander is greater than the consumer surplus for the larger consumer surplus demander. If so, then the optimal entry fee *if marginal cost is the use fee* is the consumer surplus of the lower consumer surplus demander. If not, then the optimal entry fee *if marginal cost is the use fee* is the consumer surplus of the higher consumer surplus consumer. However, it is possible that a use fee (P^*) greater than marginal cost will lead to the highest profit.

The profit margin under such a use fee is $P^* - c$. The quantity consumed in market i at price P^* is

$$Q_i = (a_i - P^*) / b_i$$

and therefore the total amount consumed by both consumers at price P^* is

$$Q = [a_1b_2 + a_2b_1 - (b_1 + b_2)P^*] / b_1b_2$$

yielding profit from the use fee of

$$[(a_1b_2 + a_2b_1)(P^* - c) - (b_1 + b_2)P^{*2} + c(b_1 + b_2)P^*]/b_1b_2$$

that is, $(P^* - c)Q$.

The consumer surplus resulting from the choice of P^* as the use fee would be

$$CS_i = (a_i - P^*)^2/2b_i$$

If double the resulting consumer surplus for the demander with the lower consumer surplus exceeds the resulting consumer surplus of the demander with the higher consumer surplus, managers will choose the smaller consumer surplus as the entry fee. How can we determine which consumer surplus is relevant?

Suppose $a_1 > a_2$. We can portray the decision by managers when seeking the best use price greater than marginal cost as dividing the range of $a_2 > P_{\text{Use}} > MC$ into regions to determine the relevant entry fee to charge. a_2 is an upper boundary because demander 2 will not participate in the market if the use fee is higher than a_2 . MC is a lower boundary because managers will never produce unless the price they can charge for output at least covers marginal cost. When a P_{Use} exists such that $CS_1 = CS_2$, either consumer surplus is the relevant entry fee. Equating the two consumer surplus values gives

$$P = [(a_1b_2 - a_2b_1) \pm \sqrt{(a_1 - a_2)^2b_1b_2}] - (a_1 - a_2)^2b_1b_2^{0.5}/(b_2 - b_1)$$

where $A = a_1b_2 - a_2b_1$, $B = (a_1 - a_2)(b_1b_2)^{0.5}$, and $C = b_2 - b_1$.

The relationship between CS_1 and CS_2 as a function of use fee (P^*) appears in Figure 9.8 for the case where $(A + B)/C < (A - B)/C$. (This is the case examined in the first example here.)

Between MC and $(A + B)/C$, $CS_2 > CS_1$, so the consumer surplus of consumer 1 (if the optimal use price is in this range) will be charged to both demanders. The entry price will thus be

$$P_{\text{Entry}} = (a_1 - P^*)^2/2b_1$$

and the profit from the entry fee will be

$$\pi_{\text{Entry}} = (a_1 - P^*)^2/b_1$$

because both demanders will pay it.

The profit earned is

$$\begin{aligned} \pi &= (a_1b_2 + a_2b_1)(P^* - c) - (b_1 + b_2)P^{*2} + c(b_1 + b_2)P^* \\ &\quad + (a_1^2b_2 - 2a_1b_2P^* + b_2P^{*2})/b_1b_2 \\ \Delta\pi/\Delta P^* &= [(a_1b_2 + a_2b_1) - 2(b_1 + b_2)P^* + c(b_1 + b_2) - 2a_1b_2 \\ &\quad + 2b_2P^*]/b_1b_2 = 0 \end{aligned}$$

Solving for P^* gives

$$P^* = [a_2b_1 - a_1b_2 + c(b_1 + b_2)]/2b_1$$

It must be checked that the resulting P^* lies in the range of

$$MC < P^* < (A + B)/C$$

Between $(A + B)/C$ and a_2 , $CS_1 > CS_2$, so the consumer surplus of demander 2 (if the optimal use price is in this range) will be charged to both demanders.

QUANT OPTION

To set the record straight, profit is maximized when

$$\begin{aligned} d\pi/dP^* &= [(a_1b_2 + a_2b_1) - 2(b_1 + b_2)P^* + c(b_1 + b_2) - 2a_1b_2 + 2b_2P^*]/ \\ & \quad b_1b_2 = 0 \end{aligned}$$

The optimal use price in this range will be

$$P^* = [a_1b_2 - a_2b_1 + c(b_1 + b_2)]/2b_2$$

Of course we must check that the resulting P^* lies in the range of

$$(A + B)/C < P^* < a_2$$

Consider the following example: $P_1 = 9 - 3Q_1$, $P_2 = 8 - 2Q_2$, and $MC = 2$. Equating consumer surpluses gives

$$(A + B)/C = 3.55 \text{ and } (A - B)/C = 8.45$$

that is, $(-6 + 2.45)/-1 = 6 - 2.45 = 3.55 < (-6 - 2.45)/-1 = 8.45$. So this meets the conditions graphed in Figure 9.8.

Viewing potential use fees where $MC < P^* < (A + B)/C$ gives

$$P^* = (24 - 18 + 10)/6 = 2.67$$

Because $MC = 2$ and $(A + B)/C = 3.55$, the price meets the constraints.

Viewing potential use fees where $(A + B)/C < P^* < a_2$ gives

$$P^* = (18 - 24 + 10)/4 = 1$$

Because $(A + B)/C = 3.55$ and $a_2 = 8$, this price does not satisfy the constraints.

Substituting $P^* = 2.67$ into the profit function here gives a profit of 16.5556.

We must still compare this result with pricing the use fee at marginal cost and the

FIGURE 9.8

The Relationship between the Demanders' Consumer Surpluses and the Use Price Chosen

The difference in the consumer surplus of demander 2 and demander 1 varies with the use price. In this example, as the use price increases above marginal cost, demander 1's consumer surplus gets closer to demander 2's consumer surplus until at use price $(A + B)/C$, the two consumer surpluses are equal. As the use price rises above $(A + B)/C$, demander 1's consumer surplus exceeds demander 2's consumer surplus. At use price a_2 until use price a_1 , only demander 1 has consumer surplus.

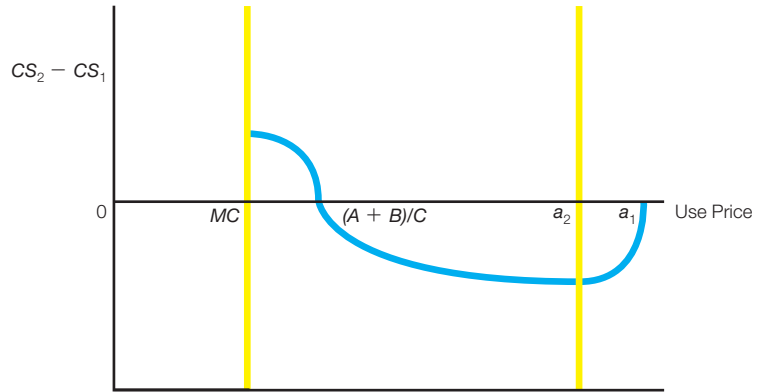


TABLE 9.1

Various Use Price and Entry Price Combinations and Their Resulting Profit

P_{Use}	CS_1	CS_2	π_{Use}	π_{Entry}	P_{Entry}	Consumer Determining Entry Fee	π
2	8.1667	9	0	16.3333	8.1667	1	16.3333
2.5	7.0417	7.5625	2.4583	14.0833	7.0417	1	16.5417
2.67	6.685	7.1111	3.185	13.37	6.685	1	16.5556
3	6	6.25	4.5	12	6	1	16.5
3.5	5.0417	5.0625	6.125	10.0833	5.0417	1	16.2083
3.55051	4.9495	4.9495	6.2667	9.899	4.9495	Either	16.165
4	4.1667	4	7.3333	8	4	2	15.3333
5	2.6667	2.25	8.5	4.5	2.25	2	13
6	1.5	1	8	2	1	2	10
7	0.6667	0.25	5.8333	0.5	0.25	2	6.3333
7.5	0.375	0.0625	4.125	0.125	0.0625	2	4.25
7.93725	0.1882	0.001	2.2896	0.002	0.001	2	2.2915
7.93725	0.1882	0.001	2.0133	0.1882	0.1882	1*	2.2915
7.95	0.1838	0.0006	2.0825	0.1838	0.1838	1*	2.2663

* In these cases only consumer 1 is served.

entry fee equal to either the larger consumer surplus (if it is more than double the smaller consumer surplus) or the smaller consumer surplus (if when doubled, it is more than the larger consumer surplus). As shown in Table 9.1, the best policy *if the firm sets the use fee equal to marginal cost* is to sell to both consumers at an entry fee of 8.1667 yielding a profit of 16.3333. Thus a use fee greater than marginal cost yields the highest profit.

But suppose the producer experimented with use fees (and charged the best resulting entry fee). The results appear in Table 9.1.

As we can see, as price rises above marginal cost, both consumers consume the product with consumer 1's consumer surplus equal to the entry fee. At a use price of 3.55051, the consumer surpluses of each consumer are the same. At use prices higher than 3.55051, both consumers consume the product with consumer 2's consumer surplus equal to the entry fee until the use price reaches 7.93725. At that price, the profit from selling to both at an entry fee equal to consumer 2's consumer surplus equals the profit from selling to just consumer 1 with an entry fee equal to consumer 1's consumer surplus. Obviously this would be true at prices between 8 and 9, but it's best to exclude consumer 2 before the use fee reaches consumer 2's maximum reservation price.

Consider a second example where the second demander's demand is $P_2 = 6 - 2Q_2$. Then $(A + B)/C = -7.35$, which makes no sense, and $(A - B)/C = 7.35$, which exceeds a_2 . In this case the graph looks like Figure 9.9.

FIGURE 9.9

A Different Relationship between the Demanders' Consumer Surpluses and the Use Price Chosen

In this case there is no user fee where the consumer surpluses are the same. Here we get a "corner solution," where the optimal use fee is at marginal cost.

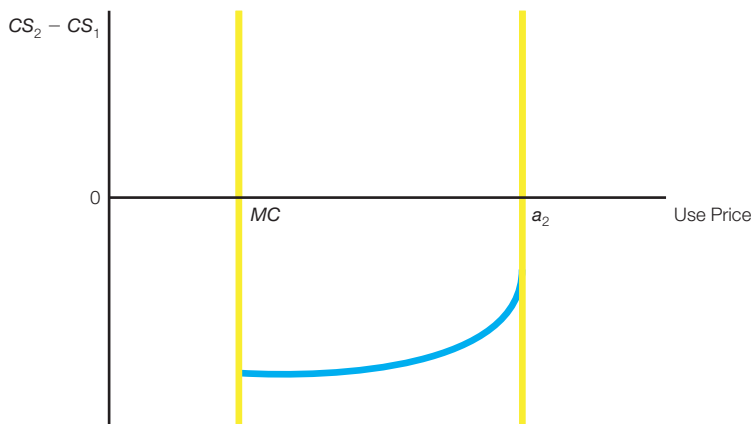


TABLE 9.2

Various Use Price and Entry Price Combinations and Their Resulting Profit

P_{Use}	CS_1	CS_2	π_{Use}	π_{Entry}	P_{Entry}	Consumer Determining Entry Fee	π
2	8.1667	4	0	8.1667	8.1667	1*	8.1667
2.67	6.685	2.7778	1.407	6.685	6.685	1*	8.09
2.95	6.1004	2.2494	1.9158	6.1004	6.1004	1*	8.0163
3	6	2.25	3.5	4.5	4	Either	8
3	6	2.25	2	6	6	1*	8
3.05	5.9004	2.1756	2.0825	5.9004	5.9004	1*	7.9829
4	4.1667	1	3.333	4.1667	4.1667	1*	7.3333
5	2.6667	0.25	4	2.6667	2.6667	1*	6.6667
6	1.5	0	4	1.5	1.5	1*	5.5

* In these cases only consumer 1 is served.

The supporting figures for Figure 9.9 are shown in Table 9.2.

In this case the best solution is to set the use price equal to marginal cost and the entry fee equal to the resulting consumer surplus of consumer 1 and not serve consumer 2. The weaker demand of demander 2 significantly constrains the profitability of the firm.

BUNDLING AND INTRAFIRM PRICING

Another sophisticated pricing strategy used by managers is bundling. **Simple bundling** occurs when managers offer several products or services as one package so consumers do not have an option to purchase package components separately. An example of simple bundling is the inclusion of a service contract with a product. **Mixed bundling** allows consumers to purchase package components either as a single unit or separately. The bundle price is generally less than the sum of the components. Season tickets to sporting events or the value meals of McDonald's are examples of a mixed bundling strategy.

Bundling is best used when there is wide variance in consumers' price sensitivity of demand and when market conditions make it difficult to price discriminate. Managers increase profit by leveraging the different valuations (reservation prices) consumers have for a product. Managers prefer to form bundles so as to create negative correlations across consumers. **Negative correlation** exists when some customers have higher reservation prices for one item in the bundle but lower reservation prices for another item in the bundle, whereas another group of customers has the reverse preferences. By bundling together the products, managers (under certain conditions) can make a greater profit than by selling the products separately.

Consumers often encounter bundling when managers offer distinct goods or services together at a packaged price. Economists distinguish between simple (pure) bundling, when the goods are offered only in the package, and mixed bundling,

LEARNING OBJECTIVES

- The Mechanics of Bundling
- When to Unbundle
- Bundling as a Preemptive Entry Strategy
- Tying at IBM, Xerox, and Microsoft
- Transfer Pricing
- Transfer Pricing: A Perfectly Competitive Market for the Upstream Product
- The Global Use of Transfer Pricing
- Summary
- Problems
- Excel Exercise: Transfer Pricing

Simple bundling When managers offer several products or services as one package so consumers do not have an option to purchase package components separately.

Mixed bundling Allows consumers to purchase package components either as a single unit or separately.

Negative correlation When some customers have higher reservation prices for one item in the bundle but lower reservation prices for another item in the bundle, whereas another group of customers has the reverse preferences.

when the goods have a packaged price as well as stand-alone prices. Examples abound. Entertainers as diverse as professional sports teams and opera companies commonly practice mixed bundling, offering tickets to individual performances as well as season tickets, subscriptions of tickets to multiple (but not all) performances, and other bundles. A record company bundles 12 songs of your favorite recording artist on a CD, and you must buy the CD (the bundle) to obtain the several songs you want (and hence obtain more songs that you do not want as much, if at all). This product is a pure bundle.

THE MECHANICS OF BUNDLING

Why do managers commonly use bundling? One reason is that it can increase the seller's profit if customers have varied tastes. In addition, it is a way to emulate perfect price discrimination when perfect price discrimination is not possible (because knowing individual reservation prices is either too difficult or expensive to pursue) or it is not legal to charge multiple prices for the same product. With a bundle, we need not know the reservation price of each consumer for each good (as in perfect price discrimination) but only the distribution of all consumers' reservation prices over the goods.

Managers need to consider other issues in choosing bundling schemes. For example, is the worth of the bundle to consumers the sum of their reservation prices for the separate goods in the bundle? That is, are the goods independent? This will be our assumption. However, we can easily envision cases of complementarity of the goods, where the goods as a bundle have a value greater than the sum of the separate reservation prices—such as software and hardware. We can also envision cases of goods, where the goods as a bundle have a value less than the sum of their separate reservation prices. And what about demand? We will assume consumers purchase no more than one unit of any good either separately or in a bundle. From the cost side, managers need to consider whether there are economies of scope or scale in the production of two goods. We will assume production costs are the same regardless of whether the goods are produced for sale separately or as a bundle; that is, the cost of a bundle is the sum of the individual costs of the two goods. A last issue is whether the goods are sold in secondary markets. We will assume that consumers do not resell the goods.

Of course managers can always sell the goods as separate items. If managers cannot price discriminate but must charge a single price for each good, we'll assume that price is the simple monopoly profit-maximizing one.

We can investigate the three possible pricing scenarios in the following three figures. Figure 10.1 shows the separate price strategy. Managers choose the optimal simple monopoly prices for good 1 and good 2 (the ones that maximize profit). Call them $p_1^{\#}$ and $p_2^{\#}$. Figure 10.1 shows the resulting consumption behavior of consumers depending on their reservation prices. For example, consumers

STRATEGY SESSION: Bundling Carbon Credits with Gas Sales

Gazprom is a Russian energy firm that sells natural gas to Europe and has profited handsomely. But now it has figured out how to make even more profit. A London subsidiary (Gazprom Marketing and Trading) has invested in a Brazilian biomass power plant (Pro-power do Brasil) that earns Gazprom Marketing and Trading carbon dioxide emission credits (because biomass uses renewable resources and is “carbon neutral” because it emits as much carbon when used as fuel as was absorbed when it was a crop). Gazprom will then bundle those credits with natural gas and sell the combination as a single product to electric utilities in Europe. The utilities need these credits to burn natural gas.

The ability to make this bundle has come about because of the 1997 Kyoto Protocol on climate change, which gave Russia a massive amount of these credits, and a 2005 European Union program of “cap and trade” that allows “dirty” utilities (heavy polluters) to purchase carbon dioxide credits that permit them to pollute and “clean” utilities (low polluters) to sell pollution credits they don’t need. Although Russia

is one of the largest producers of greenhouse gases because of antiquated technology, relatively small investments will enable them to significantly reduce emissions. This gives Gazprom access to a large amount of Russian pollution credits (the Brazilian pollution credits are just Gazprom’s test of the bundled market). To facilitate the activity, Gazprom’s banking subsidiary has established a carbon trading unit (Carbon Trade and Finance) with Dresdner Bank. Carbon Trade and Finance will then invest in modern efficient equipment at Russian polluting firms and thus earn the credits that Gazprom will use to bundle with its natural gas.

It’s estimated that the value of carbon credits in Russia is in the range of \$40 to \$60 billion. The fly in the ointment? If the Kyoto Protocol is not extended, the carbon credits could be worth nothing.

Source: Andrew E. Kramer, “Russian Energy Giant to Bundle Carbon Credits with Gas Sales,” *New York Times*, April 25, 2007, at www.nytimes.com/2007/04/25/business/worldbusiness/25carbon.html.

in the upper right cell buy both goods given their high reservation prices for the goods.

Figure 10.2 shows the strategy of pure bundling. Here managers choose the optimal pure bundle price (the one that maximizes profit). Call it $p_B^\#$. Consumers located to the right of the $p_B^\#$ line in Figure 10.2 buy the bundled product.

Figure 10.3 shows the strategy of mixed bundling. Managers choose the optimal pure bundle price (p_B^*), the optimal separate price for good 1 (p_1^*), and the optimal separate price for good 2 (p_2^*); these prices are set to maximize profit. Figure 10.3 shows the resulting consumption behavior of consumers depending on their reservation prices. This is a good example of how managerial actions influence the behavior of consumers. By increasing the purchase options, managers can track behavior in specific sectors of the market.

The optimal solution is the *greatest profit of the profit-maximizing solutions yielded by separate pricing, pure bundling, and mixed bundling*. The manager

FIGURE 10.1

Price Separately

Whether consumers purchase goods separately depends on their reservation price for the good relative to the prices charged by the seller.

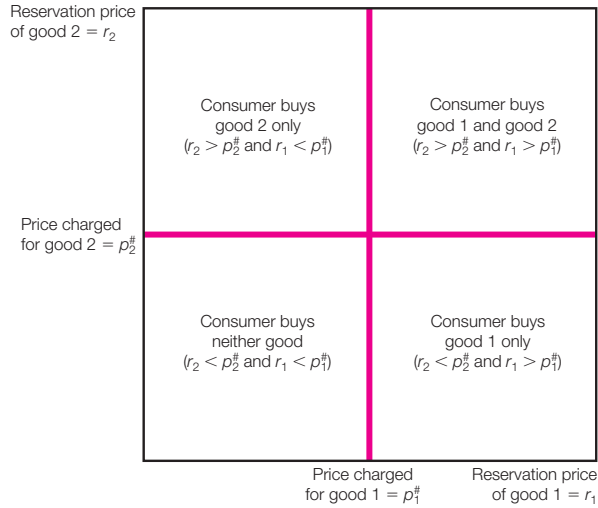
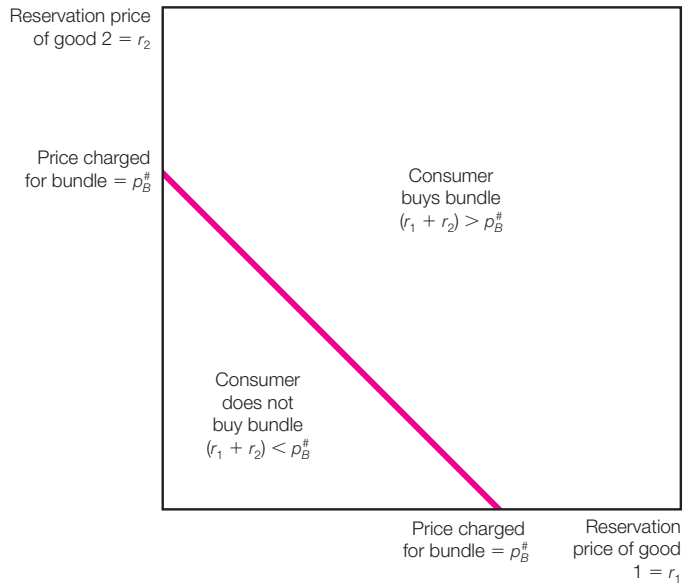


FIGURE 10.2

Pure Bundling

Whether consumers purchase the bundle depends on the sum of their reservation prices for the goods relative to the bundled price charged by the seller.



STRATEGY SESSION: Bundling College Textbooks

Sticker shock hit college students (and their parents) a number of years ago in terms of tuition rates and textbook prices, which increased at a rate more than double that of inflation of other goods and services. Textbook costs can reach over \$1,000 per year. Legislators, who have children too, have heard from their constituents and have acted.

By the way, don't blame us, your humble authors. If we had our way, textbook publishers would price where the demand elasticity was -1 where total revenue from sales of the book are maximized (because our royalties are a share of total revenue). Publishers, however, wish to maximize profits. This means charging where price equals $[\eta/(\eta+1)]MC$ where elasticity is less than -1 [$-1 > \eta$, i.e., demand is elastic], that is, a price higher than when $\eta = -1$. See, we're on your side.

Publishers have found another way to increase profits. They pure bundled the textbook with other items like a study guide, an online laboratory, and so on. Since the publishers could have mixed bundled—

that is, offered the current bundle **and** sold each item separately or just sold each item in the bundle separately and not offered the bundle at all, but rather chose only to pure bundle—pure bundling must have been most profitable and extracted the most consumer surplus from the consumer.

A new federal rule that went into effect in July 2010 now requires publishers to offer the mixed bundle. Some students may opt for the bundle for its convenience. You should never assemble your own bundle by buying each of the components separately—that would cost you more—but you may not want or need all the constituent parts of the bundle, and now you won't need to buy them to get the text.

Source: Tara Siegel Bernard, "How to Find Cheaper College Textbooks," *New York Times*, August 3, 2010, at <http://bucks.blogs.nytimes.com/2010/08/03/how-to-find-cheaper-college-textbooks>.

chooses the action that maximizes profit. These figures do not show the cost of producing the goods to keep things simple. Obviously, although the figures show the buyers' intentions, their realized transactions are a subset of those shown because certain pricing actions are precluded by the manager's profit-maximizing behavior. Solving for the profit-maximizing solutions for Figures 10.1 and 10.2 is easy because of the more limited choices of consumers. Figure 10.1 considers only individual reservation prices for the goods; and only the sum of each consumer's reservation prices is a candidate price for the pure bundle in Figure 10.2. Any other prices would unnecessarily leave consumer surplus on the table (and a profit-maximizing seller always wants to convert such consumer surplus into producer surplus).

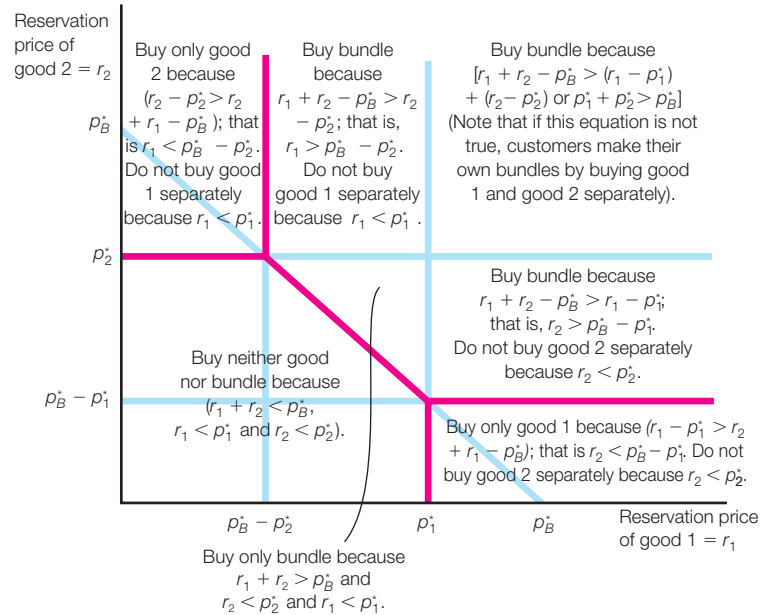
The more difficult calculation is determining the best mixed bundle prices. As we show next, the optimal prices do not have to be a reservation price of a good or a bundle. The solution is derived either by educated trial and error or via a computer program¹ that searches all separate prices and pure bundle prices and

1. Hanson and Martin, "Optimal Bundle Pricing," *Management Science* vol. 36(2), 1990, pp. 155–174, have developed such a program, as have the authors.

FIGURE 10.3

Mixed Bundling

Whether the consumer purchases the goods separately or as a bundle depends on the consumer surplus (the difference between consumers' reservation prices, or sum of their reservation prices, and the price charged by the seller). Consumers choose the goods or bundles that maximize their consumer surplus.



chooses the combination yielding the highest profit. Trial and error can be done in simple cases with few consumers and goods. Cases with many consumers and goods require a computer program. Critical to mixed bundling is creating a credible mixed bundle. **Credibility of the bundle** means that managers correctly anticipate which customers will purchase the bundle or the goods separately.

The following example shows the three types of bundling strategies when reservation prices of consumers are perfectly negatively correlated; that is, all the reservation prices lie on a line of slope -1 in the price space. Note that while the customers have a negative correlation in their reservation prices for the two goods, they exhibit *no* variation in their valuation of the bundle: They all value the bundle at \$100. The consumer reservation prices are shown in Table 10.1, and the situation is depicted in Figure 10.4.

Suppose the constant unit cost of production of each good is 1. The separate price, pure bundling, and mixed bundling cases are shown in Tables 10.2, 10.3, and 10.4.

Credibility of the bundle When managers correctly anticipate which customers will purchase the bundle or the goods separately.

TABLE 10.1

Reservation Prices of Good 1 and Good 2 of Consumers *A*, *B*, *C*, and *D*

Consumer	Reservation Price		Bundle Price
	Good 1	Good 2	
<i>A</i>	90	10	100
<i>B</i>	60	40	100
<i>C</i>	50	50	100
<i>D</i>	10	90	100

FIGURE 10.4

Example of Perfect Negative Correlation of Consumers' Reservation Prices

Consumers *A*, *B*, *C*, and *D* value each good differently, but all value the bundle of the two goods at \$100.

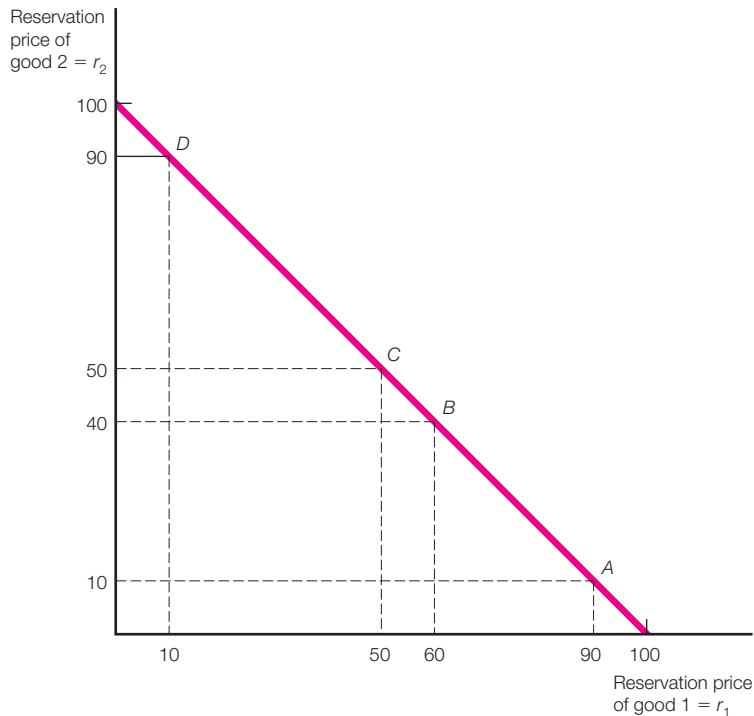


TABLE 10.2

Optimal Separate Prices for Good 1 and Good 2

Consumer	Price 1	Cost/unit	Profit/unit	Number of units	Profit
A	90	1	89	1	89
B	60	1	59	2	118
C	50	1	49	3	147
D	10	1	9	4	36

Consumer	Price 2	Cost/unit	Profit/unit	Number of units	Profit
A	10	1	9	4	36
B	40	1	39	3	117
C	50	1	49	2	98
D	90	1	89	1	89

The profit from the best separate price strategy of $P_1 = \$50$ and $P_2 = \$40$ is \$264.

TABLE 10.3

Optimal Pure Bundle Price for Consumers A, B, C, and D

Consumer	Bundle Price	Cost/Bundle	Profit/Bundle	Number of Bundles	Profit
A, B, C, D	100	2	98	4	392

The profit from the best pure bundling strategy of $P_{\text{Bundle}} = \$100$ is \$392.

Managers can always come up with a mixed bundle by pricing the individual goods at prices at which no consumer purchases the good. In some cases, *but not this one*, it is possible to increase profit through mixed bundling. Mixed bundling, therefore, always *weakly dominates* pure bundling.

If we look at mixed bundles where customers actually consume the bundle and at least one of the goods is sold separately, the pricing strategy in Table 10.5 is the best mixed bundle.

Note that consumer A does not consume the bundle because at price \$100, she receives no consumer surplus. However, consumer A consumes good 1 at \$89.99 because she receives a positive consumer surplus (of \$0.01). Likewise, consumer D does not consume the bundle because at price \$100, he receives no consumer surplus. However, consumer D consumes good 2 at \$89.99 because he receives a positive consumer surplus (of \$0.01). We discuss this logic further in another example.

TABLE 10.4

Optimal Mixed Bundle Prices

Consumer	Bundle Price	Cost/Bundle	Profit/Bundle	Number of Bundles	Profit
<i>A, B, C, D</i>	100	2	98	4	392

Consumer	Price 1	Cost/Unit	Profit/Unit	Number of Units	Profit
None	90.01	1	89.01	0	0

Consumer	Price 2	Cost/Unit	Profit/Unit	Number of Units	Profit
None	90.01	1	89.01	0	0

The profit from the best mixed bundling strategy of $P_{\text{Bundle}} = \$100$, $P_1 > \$90$, and $P_2 > \$90$ is \$392.

TABLE 10.5

Optimal Mixed Bundle Prices When Consumers Buy Bundle and at Least One of the Separately Priced Goods

Consumer	Bundle Price	Cost/Bundle	Profit/Bundle	Number of Bundles	Profit
<i>B, C</i>	100	2	98	2	196

Consumer	Price 1	Cost/Unit	Profit/Unit	Number of Units	Profit
<i>A</i>	89.99	1	88.99	1	88.99

Consumer	Price 2	Cost/Unit	Profit/Unit	Number of Units	Profit
<i>D</i>	89.99	1	88.99	1	88.99

The profit from the best mixed bundling strategy, where customers actually purchase the bundle and purchase at least one of the separately priced goods of $P_{\text{Bundle}} = \$100$, $P_1 = \$89.99$, and $P_2 = \$89.99$, is \$379.98.

Extraction When the manager extracts the entire consumer surplus from each customer.

Exclusion When the manager does not sell a good to a customer who values the good at less than the cost of producing it.

Inclusion When a manager sells a good to a consumer who values the good at greater than the seller's cost of producing the good.

In the preceding case, where pure bundling is the best pricing strategy, perfect price discrimination is completely replicated as the manager extracts *all* the consumer surplus from each customer. This goal is called **extraction**. In addition, a manager can also practice **exclusion**: not selling a good to a customer who values the good at less than the cost of producing it. Finally, a manager may want to practice **inclusion**: selling a good to a customer who values the good at greater than the seller's cost of producing the good. Perfect price discrimination extracts all available consumer surplus, does not sell to anyone for less than cost, and sells to everyone who values the good more than cost. Thus perfect price discrimination satisfies all three of the concepts defined here. These three pricing mechanisms (price separately, pure bundling, and mixed bundling) can be compared to perfect price discrimination on the three dimensions of extraction, exclusion, and inclusion.

Pricing separately should always entail exclusion; but because of its single price per good nature, it will not fulfill complete extraction or inclusion (for negatively sloped demand curves). Pure bundling can allow complete extraction (as in the preceding case); but when the sum of all demanders' reservation prices for goods does not lie on a line with a slope of -1 (that is, there is less than perfect negative correlation of reservation prices), extraction is less than complete. Mixed bundling falls someplace between pricing separately and pure bundling.

Pure bundling can also fail inclusion and exclusion. Note how the best price strategy changes when the cost of producing the goods changes. Consider the example just used but with the production cost of each good now at \$11 each. Tables 10.6, 10.7, and 10.8 show the solutions for pricing separately, pure bundling, and mixed bundling.

TABLE 10.6

Optimal Separate Prices for Good 1 and Good 2

Consumer	Price 1	Cost/Unit	Profit/Unit	Number of Units	Profit
A	90	11	79	1	79
B	60	11	49	2	98
C	50	11	39	3	117
D	10	11	-1	4	-4

Consumer	Price 2	Cost/Unit	Profit/Unit	Number of Units	Profit
A	10	11	-1	4	-4
B	40	11	29	3	87
C	50	11	39	2	78
D	90	11	79	1	79

The profit from the best separate price strategy of $P_1 = \$50$ and $P_2 = \$40$ is \$204.

TABLE 10.7

Optimal Pure Bundle Prices for Consumers A, B, C, and D

Consumer	Bundle Price	Cost/Bundle	Profit/Bundle	Number of Bundles	Profit
A, B, C, D	100	22	78	4	312

The profit from the best pure bundling strategy of $P_{\text{Bundle}} = \$100$ is \$312.

TABLE 10.8

Optimal Mixed Bundle Prices

Consumer	Bundle Price	Cost/Bundle	Profit/Bundle	Number of Bundles	Profit
A, B, C, D	100	22	78	4	312

Consumer	Price 1	Cost/Unit	Profit/Unit	Number of Units	Profit
None	90.01	11	79.01	0	0

Consumer	Price 2	Cost/Unit	Profit/Unit	Number of Units	Profit
None	90.01	11	79.01	0	0

The profit from the best mixed bundle strategy of $P_{\text{Bundle}} = \$100$, $P_1 > 90$, and $P_2 > 90$ is \$312.

However, if we look at mixed bundles where customers actually consume the bundle and at least one of the goods sold separately, Table 10.9 shows the best mixed bundle.

In this case the concept of exclusion dominates the concept of extraction. The pure bundle price of \$100 completely extracts all consumer surplus. However, the seller sells (in the bundle) good 2 to consumer A, and A values the good at only \$10, whereas it costs the seller \$11 to produce good 2. Likewise, the seller sells good 1 to consumer D in the bundle, and D values the good at only \$10, whereas it costs the seller \$11 to produce good 1. It is better for the seller to exclude consumer A from buying good 2 and consumer D from buying good 1. The seller can do that by practicing mixed bundling. The seller sacrifices \$10.01 in consumer revenue from each of consumers A and D by switching those consumers from a price of \$100 for the bundle to a price of \$89.99 for the separate goods (a total of \$20.02). But the seller saves \$22 in cost by not producing one unit of good 1 and one

TABLE 10.9

Optimal Mixed Bundle Prices When Consumers Buy Bundle and at Least One of the Separately Priced Goods

Consumer	Bundle Price	Cost/Bundle	Profit/Bundle	Number of Bundles	Profit
<i>B, C</i>	100	22	78	2	156

Consumer	Price 1	Cost/Unit	Profit/Unit	Number of Units	Profit
<i>A</i>	89.99	11	78.99	1	78.99

Consumer	Price 2	Cost/Unit	Profit/Unit	Number of Units	Profit
<i>D</i>	89.99	11	78.99	1	78.99

The profit from the best mixed bundle strategy, where customers actually purchase the bundle and at least one of the separately priced goods of $P_{\text{Bundle}} = \$100$, $P_1 = \$89.99$, and $P_2 = \$89.99$, is \$313.98. Therefore, mixed bundling is the best pricing strategy for the seller.

unit of good 2. This \$1.98 difference is the difference in profits between the best pure bundle profit of \$312 and the best mixed bundle profit of \$313.98. Inclusion and exclusion are practiced perfectly in this case of mixed bundling, but complete extraction is not. In general, optimal pricing solutions among these three methods entail a trade-off among the concepts of extraction, exclusion, and inclusion.

Suppose further that the cost of producing each good is now \$55. Tables 10.10, 10.11, and 10.12 show the solutions for pricing separately, pure bundling, and mixed bundling.

The only reason that mixed bundling yields the same profit as separate pricing here is that a bundle price is selected so that no consumer will choose the bundle—that is, a price over \$100. In cases where separate pricing is best, we can always price the bundle at a price at which no one will consume it. Therefore, mixed bundling weakly dominates pricing separately. Previously we established that mixed bundling weakly dominates pure bundling; technically mixed bundling should be a part of any bundling strategy because the profit from it is always better than or equal to that of pricing separately or pure bundling.

Although pure bundling perfectly extracts all consumer surplus in this perfectly negatively correlated reservation price example when the unit production cost is \$55, it fails miserably on exclusion. Many units (five) are sold to customers who value the good at less than its cost of production. Mixed bundling, except at a price that excludes everyone from buying the pure bundle, can do no better than

TABLE 10.10

Optimal Separate Prices for Good 1 and Good 2

Consumer	Price 1	Cost/Unit	Profit/Unit	Number of Units	Profit
A	90	55	35	1	35
B	60	55	5	2	10
C	50	55	−5	3	−15
D	10	55	−45	4	−180

Consumer	Price 2	Cost/Unit	Profit/Unit	Number of Units	Profit
A	10	55	−45	4	−180
B	40	55	−15	3	−45
C	50	55	−5	2	−10
D	90	55	35	1	35

The profit from the best separate price strategy of $P_1 = \$90$ and $P_2 = \$90$ is \$70.

TABLE 10.11

Optimal Pure Bundle Prices for Consumers A, B, C, and D

Consumer	Bundle Price	Cost/Bundle	Profit/Bundle	Number of Bundles	Profit
A, B, C, D	100	110	−10	4	−40

The best pure bundling strategy is any bundle price over \$100. No one will buy the bundle, and the profit is \$0.

TABLE 10.12

Optimal Mixed Bundle Prices at Any Pure Bundle Price over \$100
(So No Bundle Is Purchased)

Consumer	Price 1	Cost/Unit	Profit/Unit	Number of Units	Profit
A	90	55	35	1	35

Consumer	Price 2	Cost/Unit	Profit/Unit	Number of Units	Profit
B	90	55	35	1	35

The profit from the best mixed bundle strategy of $P_{\text{Bundle}} > \$100$, $P_1 = \$90$, and $P_2 = \$90$ is \$70.

2. The following examples are from “Bundling: Teaching Note,” Harvard Business School, 5-795-168, rev. July 22, 1998.

3. Call the optimal price of good 1 x . Because of the uniform distribution, this also is the optimal price of good 2. When x is chosen, it determines how many customers consume each good. Consider choosing x on the horizontal axis. Everyone to the left of x does not consume good 1, and everyone at and to the right of x consumes good 1. Consider choosing x on the vertical axis. Everyone below x does not consume good 2, and everyone at and above x consumes good 2. The total area of Figure 10.5 is $100 \times 100 = 10,000$. Viewing each of the four areas created, we can calculate the percentage of the total area occupied by each of the subareas. For instance, area A occupies $x^2/10,000$ amount of the total area. Area B occupies $(100 - x)x/10,000 = (100x - x^2)/10,000$ of the total area (as does area D). Area C occupies $(100 - x)(100 - x)/10,000 = (10,000 - 200x + x^2)/10,000$ of the total area. The number of customers in each area is the percentage of the total area times 10,000. The revenue from each area is the number of customers times the price they pay. Therefore, area A generates $0x^2$ in revenue, area B generates $x(100x - x^2)$ in revenue, area C generates $2x(10,000 - 200x + x^2)$ in revenue, and area D generates $x(100x - x^2)$ in revenue, yielding a total revenue (TR) of $20,000x - 200x^2$. We maximize TR by setting $dTR/dx = 20,000 - 400x = 0$ or $x = \$50$. Total profit is \$500,000.

pricing separately. Pricing separately extracts much of the *profitable* consumer surplus, excludes the right consumers, but does not include consumer *B* who values good 1 at \$60 (the cost of production is \$55).

Negative correlation of reservation prices enables a manager to fully extract all consumer surplus with a pure bundle when the cost of production is low. If we increase the production cost while keeping the reservation prices with perfectly negative correlation, initially mixed bundling is the profit-maximizing action; if production costs keep increasing, eventually separate pricing will maximize profit.

But negative correlation is not required to make bundling the best choice. Suppose customers are uniformly distributed over reservation prices for good 1 from \$0 to \$100 and for good 2 from \$0 to \$100. This would be a case of zero correlation of reservation prices. There are 10,000 such consumers. To keep things simple, suppose production costs of the goods are \$0. Therefore, maximizing the total revenue is the same as maximizing profit.²

The best separate prices are $P_1 = \$50$ and $P_2 = \$50$, and the profit is \$500,000.³ This is shown in Figure 10.5.

FIGURE 10.5

Optimal Separate Prices in the Case of Uniformly Distributed Consumer Reservation Prices

The optimal separate prices when the uniform distribution of consumer reservation prices is between \$0 and \$100 for both goods are \$50 for each good. Profits are \$500,000.

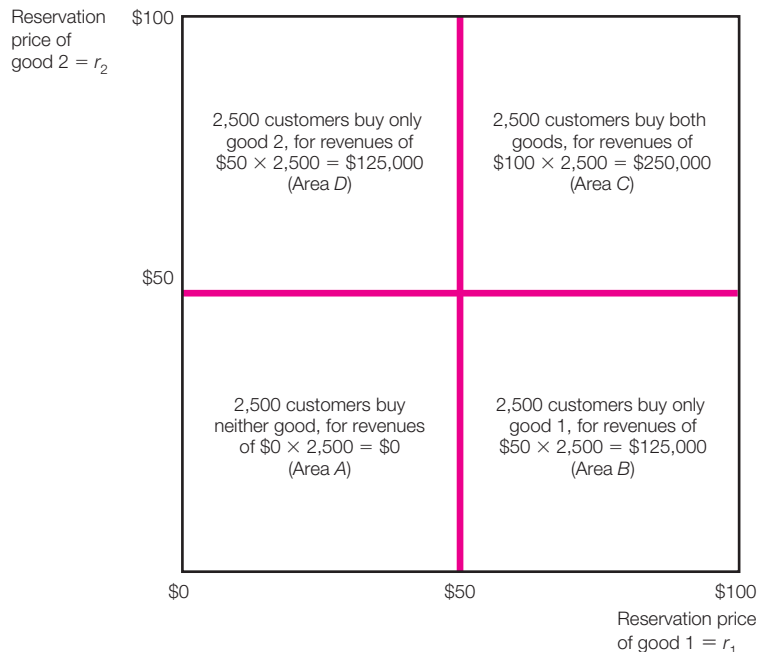
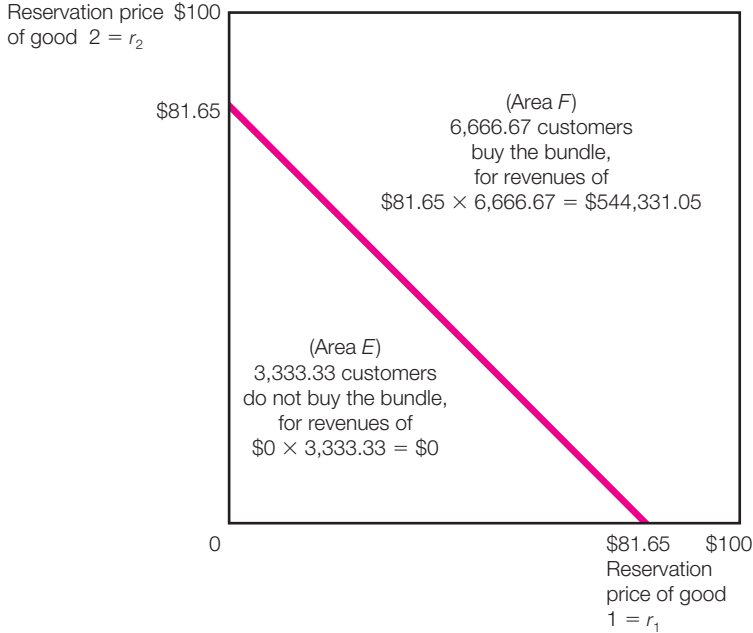


FIGURE 10.6

Optimal Pure Bundle Price in the Case of Uniformly Distributed Consumer Reservation Prices

The optimal pure bundle price when the uniform distribution of consumer reservation prices is between \$0 and \$100 for both goods is \$81.65. Profits are \$544,331.05.



The best pure bundle price is approximately \$81.65, and profit is approximately \$544,331.05.⁴ This is shown in Figure 10.6. Even without negative correlation, bundling can increase profit over simple monopoly pricing (that is, pricing separately).

The best mixed bundle has a bundle price of approximately \$86.19, $P_1 = \$66.67$, and $P_2 = \$66.67$, yielding a profit of approximately \$549,201.⁵ Therefore, mixed bundling is even better than pure bundling. This is shown in Figure 10.7. So with no negative correlation of reservation prices, the best pricing policy is mixed bundling.⁶

We may also consider quantity discounting as a form of mixed bundling. Suppose that the cost of producing the good is \$1. Table 10.13 represents consumers' reservation prices for the first unit of the good and the second unit of the good. Consumers want (at most) two units of the good. Table 10.14 shows the case of separate pricing, and Table 10.15 shows the pure bundling strategy. Finally, Table 10.16 shows the best mixed bundling strategy.

4. Call the optimal price of the bundle y . When the price of the bundle is chosen, it creates a line of slope -1 that connects the vertical axis from point y to the horizontal axis at point y . Area E has an area of $0.5(y)(y) = 0.5y^2$. Its share of the whole area is $0.5y^2/10,000$, and hence it has $0.5y^2$ customers in it. Area F has the remainder of the customers—that is, $10,000 - 0.5y^2$. The total revenue generated from area E is $\$0(0.5y^2) = \0 , and the total revenue generated from area F is $y(10,000 - 0.5y^2)$. The total revenue (TR) from pure bundling is therefore $10,000y - 0.5y^3$. Maximize TR by setting $dTR/dy = 10,000 - 1.5y^2 = 0$ or $y \approx \$81.65$. Total profit is approximately \$544,331.05.

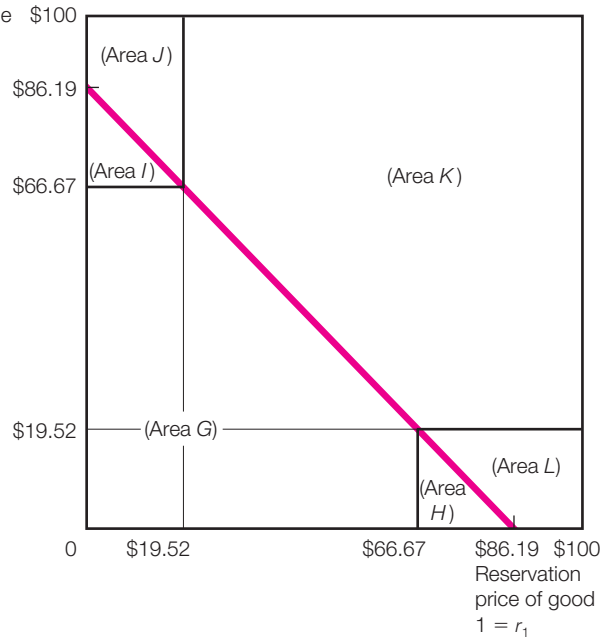
5. Call y the bundle price and x the price of good 1 (and good 2 because of the symmetry). Areas $G + H + I = 0.5y^2$. As shown in the two previous footnotes, these areas represent the number of consumers in the area. Area $H =$ Area $I = 0.5(y - x)(y - x) = 0.5y^2 - xy + 0.5x^2$. Therefore, Area $G = 2xy - 0.5y^2 - x^2$. Area $I +$ Area $J =$ Area $H +$ Area $L = (100 - x)(y - x) = 100y - 100x + x^2 - xy$. Hence, Area $J =$ Area $L = 100y - 100x + x^2 - xy - (0.5y^2 - xy + 0.5x^2) = 100y - 100x - 0.5y^2 + 0.5x^2$. Area $J +$ Area $K +$ Area $L = 10,000 - 0.5y^2$. Area $K = 10,000 - 0.5y^2 - (200y - 200x - y^2 + x^2) = 10,000 - 200y + 200x + 0.5y^2 - x^2$. Consumers in Area $I +$ Area J purchase only good 2 and yield revenues of $x(100y - 100x + x^2 - xy) = 100xy - 100x^2 + x^3 - x^2y$. Consumers in Area $H +$ Area L purchase only good 1 and yield revenues of $x(100y - 100x + x^2 - xy) = 100xy - 100x^2 + x^3 - x^2y$. Consumers in Area K purchase the bundle and yield revenues of $y(10,000 - 200y + 200x + 0.5y^2 - x^2) = 10,000y - 200y^2 + 200xy + 0.5y^3 - x^2y$. Consumers in Area G buy nothing and yield no revenues. Total revenue (TR) is therefore

(Continued)

FIGURE 10.7

Optimal Mixed Bundle Pricing in the Case of Uniformly Distributed Consumer Reservation Prices

The optimal mixed bundle pricing when the uniform distribution of consumer reservation prices is between \$0 and \$100 for both goods is $P_1 = \$66.67$, $P_2 = \$66.67$, and $P_{\text{Bundle}} = \$86.19$. Profit is \$549,201.



$10,000y - 200y^2 + 400xy + 0.5y^3 - 3x^2y - 200x^2 + 2x^3$. Total revenues are maximized where $\partial TR/\partial x = 0$ and $\partial TR/\partial y = 0$; $\partial TR/\partial x = 400y - 6xy - 400x + 6x^2 = 0$.

This yields $400x - 400y = 6x^2 - 6xy$ or $400(x - y) = 6x(x - y)$, or $6x = 400$, which yields $x = \$66.67$. $\partial TR/\partial y = 10,000 - 400y + 400x + 1.5y^2 - 3x^2 = 0$. Substituting $x = \$66.67$ into $\partial TR/\partial y$ yields $10,000 - 400y + 26,666.67 + 1.5y^2 - 13,333.33 = 1.5y^2 - 400y + 23,333.33 = 0$. Solving via the quadratic formula yields $y = \$86.19$. Substituting $y = \$86.19$ and $x = \$66.67$ into $TR = 10,000y - 200y^2 + 400xy + 0.5y^3 - 3x^2y - 200x^2 + 2x^3 = \$549,201$.

6. Bundling can also work if reservation prices are positively correlated. Consider the case where consumer A values good 1 at 11 and good 2 at 24, consumer B values good 1 at 15 and good 2 at 45, and consumer C values good 1 at 16 and good 2 at 15. The reservation prices are weakly positively correlated (0.037); that is, the slope of a linear regression is 0.214. The cost of good 1 is 5 and the cost of good 2 is 10. The best separate prices of $P_1 = 15$ and $P_2 = 45$ yield a profit of 55. The best pure bundle price is $P_B = 31$, yielding a profit of 48. But the best choice is a mixed bundle of $P_B = 60$, $P_1 = 16$, and not offering good 2 separately. Consumer B buys the bundle, consumer C buys good 1, and consumer A buys nothing. The profit is 56.

We give one final example of bundling to demonstrate how tricky calculating an optimal mixed bundle pricing package can be. The example demonstrates the point made previously about having to consider only customers' reservation prices as candidates for optimal separate prices and only the sum of customers' reservation prices as candidates for pure bundling. It also shows that in mixed bundling, the optimal prices need not be any customer's reservation price (or sum of reservation prices). And the consumer selects the good or bundle that leaves her with the greatest consumer surplus.

The complexity of solving for the optimal bundle is that managers have no $\Delta\pi/\Delta Q = 0$ formula to help derive the optimal pricing scheme. The procedure is more trial and error than derivation. The mixed bundle pricing package consists of prices for individual components and a single price for the product (or service) bundle. A manager can maximize profit even if the prices of the individual goods

TABLE 10.13

Reservation Prices for the First and Second Units of a Good by Consumers *A* and *B*

Consumer	Reservation Price of Good	
	First Unit	Second Unit
<i>A</i>	4	1.5
<i>B</i>	3.99	3

TABLE 10.14

Optimal Separate Prices for the Good

Price of Good	Cost/Unit	Profit/Unit	Number of Units	Profit
4	1	3	1	3
3.99	1	2.99	2	5.98
3	1	2	3	6
1.5	1	0.5	4	2

So the best separate price is \$3, and the profit is \$6.

TABLE 10.15

Optimal Pure Bundle Price for Two Units of the Same Good

Price of Bundle	Cost/Bundle	Profit/Bundle	Number of Bundles	Profit
5.5	2	3.5	2	7
6.99	2	4.99	1	4.99

So the best pure bundling strategy is a price of \$5.5, and the profit is \$7.

or the bundle are different from the reservation prices of consumers. This cannot be true when considering just separate pricing or pure bundling. Whether we deal with prices different than reservation prices or their sums depends on trade-offs from the consumers' view of consumer surplus and from the producers' view of producer surplus.

TABLE 10.16

Optimal Mixed Bundling Prices for the Case of a Single Good

Price of Bundle	Cost/Bundle	Profit/Bundle	Number of Bundles	Profit
6.99	2	4.99	1 (B)	4.99

Price of Good	Cost/Unit	Profit/Unit	Number of Units	Profit
4	1	3	1 (A)	3

So the best mixed bundle and the best overall pricing strategy are to price a unit of the good at \$4 and a bundle of two goods at \$6.99. This yields a profit of \$7.99.

Consider the scenario in Table 10.17. There are three consumers (or consumer groups)—*A*, *B*, and *C*, for simplicity, with an equal number of consumers in each group, each wanting no more than one of each good at their reservation price or less—and two goods, *X* and *Y*. Both products cost a constant \$4 to produce. The best separate prices are shown in boldface in Table 10.18. Therefore, the best separate pricing strategy is price *X* at \$12 and price *Y* at \$8, which yields a profit of \$16. The best pure bundling price is shown in Table 10.19 in boldface. The best pure bundle price is \$13.33, which yields a profit of \$15.99.

Note the negative relationship (correlation) among the customer's reservation prices. Note also that in considering separate prices, you need never consider a nonreservation price as a pricing candidate. For instance, suppose you investigated pricing good *X* at \$5. Two customers will purchase at that price (*A* and *B*, who have reservation prices of \$5.33 and \$12, respectively). Your profit with such a price is $(\$5 - \$4)2 = \$2$. But when you price at \$5, you leave consumer surplus on the table. Consumer *A* is willing to pay \$5.33, but you do not ask her to. As a result, you sacrifice \$0.33 in profit (not only on customer *A* but also on customer *B* if you had offered good *X* at \$5.33). By lowering your price to \$5, you pick up no additional sales and sacrifice \$0.66 in profit. If you do not charge the reservation prices of the customers, you cannot maximize profit using a separate price strategy. The same is true for pure bundling. Why shouldn't you consider pricing the bundle at \$14.50 as one of your candidates? Because at \$14.50 you get only customer *B* to buy the bundle, and she would have purchased the bundle at \$15. You would leave consumer surplus on the table.

Consider now a mixed bundling strategy with a bundle price of \$13.33, a price of good *X* at \$10.32, and a price of good *Y* at \$10.32. It first might appear that all customers would buy the bundle at \$13.33 (because *B* gets consumer surplus

TABLE 10.17

Consumer Reservation Prices for Good X and Good Y (in Dollars)

	Reservation Prices for Goods by Consumer		
	Good X	Good Y	Both X and Y
Consumer A	5.33	8	13.33
Consumer B	12	3	15
Consumer C	3	11	14

TABLE 10.18

Best Separate Price Strategy

	Cost/Unit	Profit/Unit	Number of Units	Profit
Price of X				
5.33	4	1.33	2	2.66
12.00	4	8.00	1	8.00
3.00	4	-1.00	3	-3.00
Price of Y				
8.00	4	4.00	2	8.00
3.00	4	-1.00	3	-3.00
11.00	4	7.00	1	7.00

TABLE 10.19

Best Pure Bundling Strategy

Price of Bundle	Cost/Bundle	Profit/Bundle	Number of Bundles	Profit
13.33	8	5.33	3	15.99
15.00	8	7.00	1	7.00
14.00	8	6.00	2	12.00

of $\$15 - \$13.33 = \$1.67$ and C gets consumer surplus of $\$14 - \$13.33 = \$0.67$). But if the price of X is $\$10.32$, consumer B gets a higher consumer surplus of $\$12 - \$10.32 = \$1.68 > \1.67 if she buys good X alone; and if the price of good Y is $\$10.32$, consumer C gets a higher consumer surplus of $\$11 - \$10.32 = \$0.68 > \0.67 if she buys good Y alone. But consumer B does not get to consume good Y if she does not buy the bundle, and consumer C does not get to consume good X if she does not buy the bundle—doesn't that count? Yes, but they are still better off with the larger consumer surplus from consuming just one good.

Suppose consumer B has $\$15$. If she buys the bundle for $\$13.33$, she will have consumer surplus of $\$1.67$ and both goods. But if she buys just good X for $\$10.32$, she will have good X and $\$4.68$ left over. She has $\$1.68$ in consumer surplus from good X ; and although she does not have good Y , she has $\$4.68$ in cash ($\$3$ of which is not part of the consumer surplus from good X). But $\$4.68$ in cash instead of good Y is attractive to consumer B because good Y is worth only $\$3$ to her. (She has the equivalent of good Y with the $\$3$ in cash that is not associated with the consumer surplus for good X ; recall that the definition of reservation price is that a person is exactly indifferent between a good and the amount of the reservation price.) Therefore, $\$4.68$ in cash and good X is a better position for consumer B than $\$1.67$ in cash and both goods (which are worth only $\$4.67$ to her). Analogously, consumer C , starting off with $\$14$, is better off with $\$3.68$ in cash and good Y rather than having $\$0.67$ in cash and both goods (because good X is worth $\$3$ to her).

How much profit do managers earn with such a mixed bundle? Table 10.20 demonstrates the profit improvement to $\$17.97$. The profit of $\$17.97$ dominates the profit of $\$16$ available from the best separate pricing strategy and the profit of $\$15.99$ available from the best pure bundling strategy. The secret is to see if we can pull customers out of the best pure bundling strategy and increase profit with a credible bundle. The best pure bundle yields $\$5.33$ profit per customer. We retain that net profit for customer A and pull customers B and C out at higher profit margins, thus increasing the firm's profit. In the case of separate pricing, we need to ask whether we can put some customer(s) in a bundle and do better. We sacrifice profit margin from consumer B (down from $\$8$ in the best separate price situation), but we more than make up that loss with the tremendous gain on consumers A and C , who were yielding only $\$4$ each under the best separate pricing strategy. (We are down $\$1.68$ from customer B but up $\$1.33$ from customer A and up $\$2.32$ from customer C ; so we gain $\$1.97$, the difference between the mixed bundle profit of $\$17.97$ and the best separate price profit of $\$16$.)

If we price good X at $\$10.34$, customer B does not buy it (it yields a consumer surplus of only $\$12 - \$10.34 = \$1.66$, and she can get $\$15 - \$13.33 = \$1.67$ by buying the bundle). So pricing good X at $\$10.34$ and the bundle at $\$13.33$ is

TABLE 10.20

Best Mixed Bundling Strategy

Price of Bundle	Cost/Bundle	Profit on Each	Total Number	Total Profit
13.33	8	5.33	1 (consumer A)	5.33
Price of X	Cost/Unit	Profit/Unit	Total Number	Total Profit
10.32	4	6.32	1 (consumer B)	6.32
Price of Y	Cost/Unit	Profit/Unit	Total Number	Total Profit
10.32	4	6.32	1 (consumer C)	6.32
				Sum of Profit 17.97

not a credible mixed bundle because someone you didn't want to buy the bundle (consumer *B*) does.

As you can see, mixed bundling need not charge the reservation prices of a consumer for the items. But we see a lot of mixed bundling in the market—so it is important to know that experimentation plays an important role and pure bundling and separate pricing are not always the best strategies.

WHEN TO UNBUNDLE

It is important to remember that the concept of bundling entails a null case of pricing the bundled goods (services) separately. Just because bundling is the optimal pricing strategy at time t does not mean it is the optimal pricing strategy at time $t + 1$. Managers must reassess their markets periodically to see if changed conditions warrant new prices, including an unbundling of commodities.

Table 10.21 shows the reservation prices of consumers *A*, *B*, and *C* for goods *X* and *Y* at time t . Consumers want, at most, one of each good. The unit cost of each good is 3, and a bundle of the two goods costs the producer 6. The reservation price of a consumer for a bundle of the goods is the sum of his reservation prices for the goods. The producer cannot price discriminate.

The optimal (profit-maximizing) strategy for the producer under these conditions is to price good *X* at $P_X = 8.32$, price the bundle at $P_B = 10.33$, and not offer good *Y* separately. This policy yields a profit of 13.98. The best pure bundle

TABLE 10.21

The Reservation Prices for Consumers *A*, *B*, and *C* for Good *X*, Good *Y*, and a Bundle of Good *X* and Good *Y*

	Reservation Price		
	Good <i>X</i>	Good <i>Y</i>	Bundle of Good <i>X</i> and Good <i>Y</i>
Consumer <i>A</i>	5	5.33	10.33
Consumer <i>B</i>	3	10	13
Consumer <i>C</i>	9	2	11

would be to offer the bundle at $P_B = 10.33$, yielding a profit of 12.99. The best separate price policy would be $P_X = 9$ and $P_Y = 10$, yielding a profit of 13.

Figure 10.8 shows the best mixed bundling policy versus the best pure bundle and best separate price policies and the use of the method of “crow’s feet”⁷ to solve the bundling problem.

Pure bundling at $P_B = 10.33$ gives the lowest profit of the three types of pricing at 12.99. Each consumer contributes 4.33 ($= 10.33 - 6$) to the profit. If managers want to pull consumer *B* out of the bundle by a separate price policy, they will have to offer *B* a greater consumer surplus than she receives from the bundle (she currently gets $2.67 = 13 - 10.33$) and so must offer more than 2.67 to pull consumer *B* from the bundle—this is shown as length *b* in Figure 10.8). Only good *Y* is at play here because consumer *B* is willing to pay the *MC* only for good *X* (so managers can’t reduce the price and make a profit). Because *B* will pay 10 for good *Y*, if managers reduce the price of good *Y* to $10 - 2.67^+ = 7.33^-$, consumer *B* will have consumer surplus of $10 - 7.33^- = 2.67^+$; this beats the consumer surplus from the bundle, so she will defect from the bundle. However, this means the profit from consumer *B* falls to $7.33^- - 3 = 4.33^-$, which is less than before (and without changing consumers *A* and *C*’s behavior and hence their profit contribution). Thus, managers should not want to pull consumer *B* from the bundle.

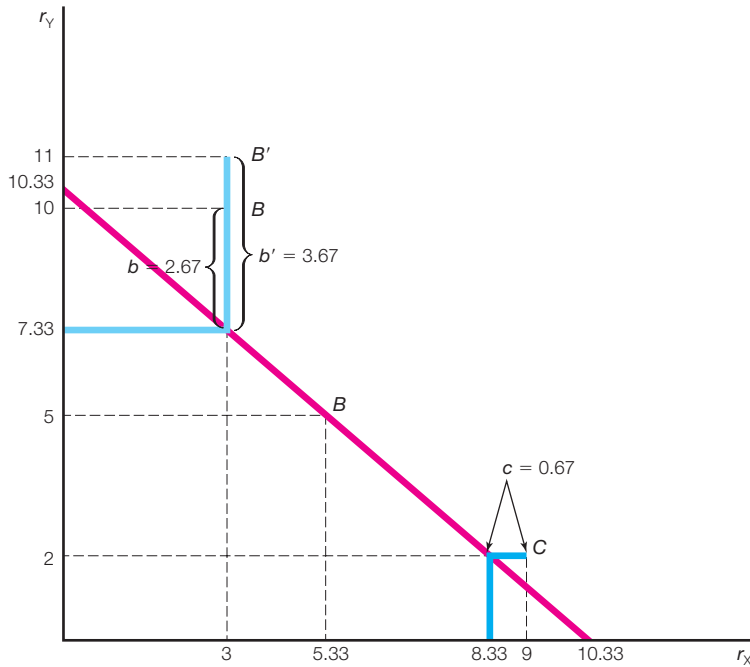
If managers want to pull consumer *C* out of the bundle by a separate price policy, they will have to offer *C* a greater consumer surplus than he receives from the bundle (he currently gets $0.67 = 11 - 10.33$) and so must offer more than 0.67 to pull consumer *C* from the bundle (this is shown as length *c* in Figure 10.8). Only good *X* is in play here because consumer *C* is not willing to pay the *MC* for good *Y*. Because *C* will pay 9 for good *X*, if managers reduce the price of good *X* to $9 - 0.67^+ = 8.33^-$, consumer *C* will have consumer surplus of $9 - 8.33^- = 0.67^+$; this beats the consumer surplus from the bundle, so he will defect from the bundle. This means the profit from consumer *C* increases to $8.33^- - 3 = 5.33^-$, which is more (by 0.99^+) than before (and without changing

7. Our former Wharton colleague Matt White introduced us to the “crow’s feet” method.

FIGURE 10.8

Depiction of Bundling Problem in Table 10.21

The bundling problem is solved by analyzing the “crow’s feet” (the large blue lines) in this figure. The “crow’s feet” method is extended when the reservation price of consumer B for good Y increases from 10 to 11.



consumers A and B 's behavior and hence their profit contribution). Thus managers do want to pull consumer C from the bundle.

Pulling consumer A from the bundle lowers profits from A and also from the other consumers. To pull A from the bundle, suppose P_X is dropped to 5^- and P_Y is dropped to 5.33^- . A will buy each good separately, but profit from A drops to $5^- - 3 + 5.33^- - 3 = 4.33^-$ (less than before). In addition, consumer B will defect from the bundle (now getting a $10 - 5.33^- = 4.67^+$ consumer surplus from good Y), and consumer C will defect from the bundle (now getting a $9 - 5^- = 4^+$ consumer surplus from good X). Profit from B falls to $5.33^- - 3 = 2.33^-$ (much less than in the bundle), and profit from C falls to $5^- - 3 = 2^-$ (much less than in the bundle). Thus the mixed bundle is much better than the pure bundle priced at $P_B = 10.33$.

How about the separate price solution versus the mixed bundle? Consumer A purchases nothing under the optimal separate price solution of $P_X = 9$ and $P_Y = 10$ and hence contributes nothing to profit. Pulling her into a bundle at

10.33 will increase profit from zero to 4.33. Consumer B buying good Y at 10 contributes 7 to profit, and consumer C buying good X at 9 contributes 6 to profit. Figure 10.8 shows that bringing consumer B into the bundle will cause B to contribute 2.67 (length b) less in profit than before ($10.33 - 6 = 4.33$ versus $10 - 3 = 7$). Pricing good X at 8.33^- will keep consumer C from joining the bundle (which he will do if the price of X remains at 9). This decreases the profit from consumer C by 0.67^+ , shown as length c in Figure 10.8 ($8.33^- - 3 = 5.33^-$ versus $9 - 3 = 6$). But on net, where does it leave total profit? The mixed bundle increases profit from A by 4.33, decreases the profit from B by 2.67, and decreases the profit from C by 0.67^+ for a net increase in profit of $4.33 - 2.67 - 0.67^+ = 0.99^-$.

Thus Figure 10.8 shows the optimality of the mixed bundle pricing policy of $P_B = 10.33$, $P_X = 8.33^-$, and not offering good Y separately over the best pure bundle strategy of $P_B = 10.33$ and the best separate price strategy of $P_X = 9$ and $P_Y = 10$. Note that the verbal explanation here is shown by the “crow’s feet” (the large red and blue lines radiating from the $P_{\text{Bundle}} = 10.33$ line in Figure 10.8).

Suppose demand conditions change. In particular, suppose consumer B ’s reservation price for good Y increases from 10 to 11 as shown in Figure 10.8 (with the consumer’s new position shown as B'). Note that B ’s reservation price for the bundle has now increased to 14 ($3 + 11$).

If consumer B buys the bundle at 10.33, she gets a consumer surplus of $3.67 (= 14 - 10.33)$ and contributes a profit of $4.33 (= 10.33 - 6)$. However, if managers do not offer the bundle, they will make $8 (= 11 - 3)$ from consumer B by charging B ’s reservation price for good Y . At the same time, charging consumer C his reservation price of 9 for good X yields a profit of $6 (= 9 - 3)$ from consumer C . Consumer A will buy neither good and hence contributes nothing to profit. However, total profit is now $14 (= 8 + 6)$, which exceeds the profit of pure bundling at $P_B = 10.33$ of 12.99 and the profit of mixed bundling at $P_B = 10.33$, $P_X = 8.33^-$, and not offering good Y separately of 13.99^- . In fact, an increase in consumer B ’s reservation price for good Y to 10.98^+ will make the unbundling policy the most profitable.

When consumer B was at point B (not B'), the separate price profit was $13.99^- - 13 = 0.99^-$ behind the optimal (mixed bundling) pricing policy. Raising the reservation price of good Y for consumer B by 1 enables managers to profitably change their pricing policy. Again, note that the “crow’s feet” tell the same story.

Here’s another case where managers will want to unbundle. If consumer C ’s good X reservation price increases to 10 (and all other consumers’ reservation prices remain the same), managers could also unbundle and sell both good X and good Y each at 10 and earn $14 (= 10 - 3 + 10 - 3)$. In fact, any combination of increases in B ’s reservation price for good Y and C ’s reservation price for good X greater than 0.98 should cause managers to unbundle the products and just sell each good separately.

Consider the new scenario where B 's reservation price for good Y is 11. Suppose a bundle price of 10.33. If consumer B buys the bundle, she receives a consumer surplus of $3.67 (= 3 + 11 - 10.33)$ and yields a profit of $4.33 (= 10.33 - 6)$. To pull B out of the bundle, she'd have to be offered a consumer surplus greater than 3.67. This would entail offering good Y at a price of less than 7.33 (say 7.33^-). Such a price would yield a profit from consumer B of less than the 4.33 available from selling B the bundle at 10.33 ($7.33^- - 3 = 4.33^-$). No other consumption would be affected by dropping the price of good Y to 7.33^- , and profit is reduced by this move.

By taking away the option of the bundle from the mixed bundling scenario and just charging $P_X = 9$ and $P_Y = 11$, Figure 10.8 shows the gain in profit. For B , who bought the bundle when it was available, profit increases from $4.33 (= 10.33 - 6)$ to $8 (= 11 - 3)$ or by $3.67 (= 8 - 4.33)$ or by b' in Figure 10.8. Individual C now buys just good X for 9, yielding a profit of $6 (= 9 - 3)$, whereas under mixed bundling, he bought just good X at 8.33^- , yielding a profit of $5.33^- (= 8.33^- - 3)$. Profit from C has thus increased by $0.67^+ (= 6 - 5.33^-)$ or by c in Figure 10.8. Without the bundle available, A consumes neither good and hence the profit of $4.33 (= 10.33 - 6)$ she contributed to the firm under the mixed bundle has been lost. But the combined increase in profit from B and C of $4.34^+ (= 3.67 + 0.67^+)$ exceeds the loss in profit from A of 4.33, yielding a net gain of $0.01^+ (= 4.34^+ - 4.33)$. This is shown by the difference in the separate price profit of 14 and the mixed bundling profit of 13.99^- .

So although some form of bundling (pure or mixed) will many times increase a firm's profit, unbundling can also increase profit. It all depends on the reservation prices and the costs of production. The diagrammatic technique shows how to measure the conditions under which managers can profitably change prices.

BUNDLING AS A PREEMPTIVE ENTRY STRATEGY

In addition to being a strategy to enhance profits, bundling is also used to deter entry by potential rivals. Suppose managers at the Alpha Company have developed a bundle made up of product W and product S , which they plan to sell for $\$X$. The question for us to answer is this: What is X ?

The Beta Company is developing a product (C) that is a close substitute to W . The Gamma Company is developing a product (N) that is a close substitute to S . Managers at both Beta and Gamma want to bring their products to market. Only Alpha has the financial ability to produce both products as a bundle. Alpha's entry cost to the market is 30. It would cost Beta 17 to enter the market, and it would cost Gamma 17 to enter the market. If Alpha managers were to produce each product separately, it would cost them 15 to enter each market. In all cases, these entry costs are independent of the number of units sold. Note that Alpha's entry costs

are 15 for each product regardless of whether the product is sold separately or is included in a bundle.

The market for the services provided by W and C , by S and N , and for a bundle that provides both services is presented in Table 10.22. Suppose consumers regard Beta's product C as comparable to Alpha's product W and regard Gamma's product N as comparable to Alpha's product S . Consumers also regard making their own bundle (buying Beta's C and buying Gamma's N) as comparable to Alpha's bundle of W and S . The numbers in the table are the consumers' reservation prices for each product. It costs each producer 2 to distribute each product (and hence it costs Alpha 4 to distribute its bundle).

For simplicity, assume there is one of each consumer type and the table represents the long-run demand for the products. Consumers want, at most, one of each product. The companies cannot form joint ventures. The consumers' reservation prices are shown in Table 10.22.

The goods are perfectly negatively correlated. Normally this suggests a pure bundle strategy. Indeed, if Alpha is the only participant in this market, the pure bundling strategy (with a bundle price of 30) yields all consumers purchasing and a net revenue exclusive of entry costs of $(30 \times 3) - (4 \times 3) = 78$. Alpha's entry costs are 30, yielding a profit of 48. This vastly exceeds the best separate price strategy of $P_W = P_S = 15$, which yields a profit of $(15 \times 4) - (2 \times 4) - (15 \times 2) = 22$, or the mixed bundling strategy of the bundle priced at 30 and $P_W = P_N = 20$, which yields a profit of $[(30 - 4) \times 1] + [(20 - 2) \times 2] - (15 \times 2) = 32$.

But if Alpha faces the entry threat from Beta and Gamma, it cannot price the bundle at 30. If it does, Beta could enter and sell C for as little as $(23/3)$ and make money. Likewise, Gamma could enter and sell N for as little as $(23/3)$ and make money. If both Beta and Gamma priced at $(23/3)$, all three consumers

TABLE 10.22

The Reservation Prices for Consumers A , B , and C for Good W or C , Good S or N , and a Bundle of Good W and Good S or a Bundle of Good C and Good N

Consumer Class	Reservation Price for Product W or C	Reservation Price for Product S or N	Reservation Price for bundle W and S or C and N
A	10	20	30
B	15	15	30
C	20	10	30

STRATEGY SESSION: How *The New Yorker* Used Bundling

The New Yorker is a wonderfully written magazine with witty, informative, entertaining articles. But apparently the market for high-quality journalism has fallen on hard times as consumers have switched their cultural and media preferences.

Magazines depend on subscriptions and newsstand sales and, most of all, advertising for their revenues. And advertising revenues depend on the number of magazines sold because advertisers pay more if the circulation to their demographic group is higher.

Over the years, circulation and advertising revenues fell. It appeared that *The New Yorker* could

not cover its costs with its revenues. Thus *The New Yorker's* publisher (Conde Nast) came up with a bundling strategy. Conde Nast also publishes *Architectural Digest* and *Vanity Fair* (which are doing well). The bundling strategy? If a company wanted to advertise in *Architectural Digest* or *Vanity Fair*, it also had to advertise in *The New Yorker*. This proved to be a profitable strategy, and *The New Yorker* is still publishing.

Source: "There's Less Buzz and Less Lunch at *The New Yorker*," *New York Times*, Monday, June 28, 1999.

would purchase *C* from Beta, giving Beta revenue of 23. Beta's entry cost is 17, and its cost of distributing *C* to the consumers would be $2 \times 3 = 6$ for a total cost of 23. The same would hold for Gamma. Both Beta and Gamma would be earning a normal profit (and zero excess profit).

If Alpha prices the bundle at slightly less than $(46/3)$, say at $(46^-/3)$, then neither Beta or Gamma can enter the market because neither can cover the $(23/3)$ cost. How will Alpha do? All three consumers will purchase Alpha's bundle at $(46^-/3)$, yielding Alpha revenue of 46^- . Alpha's cost of entry would be $15 \times 2 = 30$, and the cost of distributing the three bundles would be $4 \times 3 = 12$ for a total cost of 42. Profit would be $46^- - 42 = 4^-$. This is a far cry from 48. But it does leave Alpha as the sole producer.

The threat of entry can significantly reduce monopoly profit. But at the same time the use of bundling can preclude entry and keep a profitable market for the bundler.

Despite the lack of a rigorous analytical solution, following a few simple guidelines will help managers construct more effective bundling policies:

1. If goods' reservation prices are positively correlated, pure bundling can do no better than separate pricing (but mixed bundles might).
2. If the marginal cost of producing a good exceeds its reservation price, in general you should think carefully about selling it.⁸

8. Sometimes it may be more profitable to sell a good in a pure bundle even though the reservation price of a bundle buyer for a good in the bundle is less than the cost of producing the good. Consider a firm choosing between only separate pricing and pure bundling. The reservation prices for good 1 and good 2 for consumers A through F are (\$70, \$30), (\$80, \$20), (\$75, \$25), (\$75, \$15), (\$84, \$16), and (\$90, \$10), respectively. The unit cost of production of good 1 is \$70 and of good 2 is \$20. The optimal separate prices are $P_1 = \$80$ and P_2 either \$30 or \$25, yielding a profit of \$40. The optimal pure bundling price of \$100 yields a profit of \$50. Consumers E and F, who buy the bundle, value good 2 at less than its cost and are not excluded because that would decrease the firm's profit. If mixed bundling is allowed, the exclusion problem can be solved. The bundle would be priced at \$100, and good 1 would be priced at \$83.99. This yields a profit of \$57.98, and E and F consume only the good (good 1) they value at greater than cost.

STRATEGY SESSION: When to Unbundle

In July 2011, managers of Netflix unbundled its streaming DVD and its mail DVD plans into two separate product lines, that are priced separately. Previously, the products were only sold as a pure bundle for \$9.99/month. Now each product would cost \$7.99/month; in other words, making your own bundle cost \$15.98.

Initial reaction of Netflix customers seemed to be overwhelmingly negative, with an unscientific survey in *USA Today* claiming that 74% would switch away from Netflix. Netflix managers were stunned by the negative reaction. They claimed to have thought carefully prior to their decision.

Let's use the techniques in the text to see what must be going on to make this a viable strategy for managers. We'll concentrate just on the revenue implications first and then make some cost assumptions. The diagram on the next page shows the price separately and the pure bundling situations.

Under the pure bundling pricing, consumers in areas *B*, *C*, *D*, and *E* will buy the bundle. Assume there are *A* consumers in area *A*, *B* consumers in area *B*, and so on. Thus, total revenue from the previous pricing policy is

$$(X) \quad \$9.99B + \$9.99C + \$9.99D + \$9.99E$$

Under the price separately policy, consumers in areas *A*, *B*, and *C* will buy the streaming DVDs and consumers in areas *C*, *E*, and *F* will buy the mail DVDs. Thus, total revenue from pricing separately is

$$(Y) \quad \$7.99A + \$7.99B + \$15.98C \\ + \$7.99E + \$7.99F$$

From a revenue standpoint, Netflix managers chose the correct strategy if $(Y) > (X)$ or

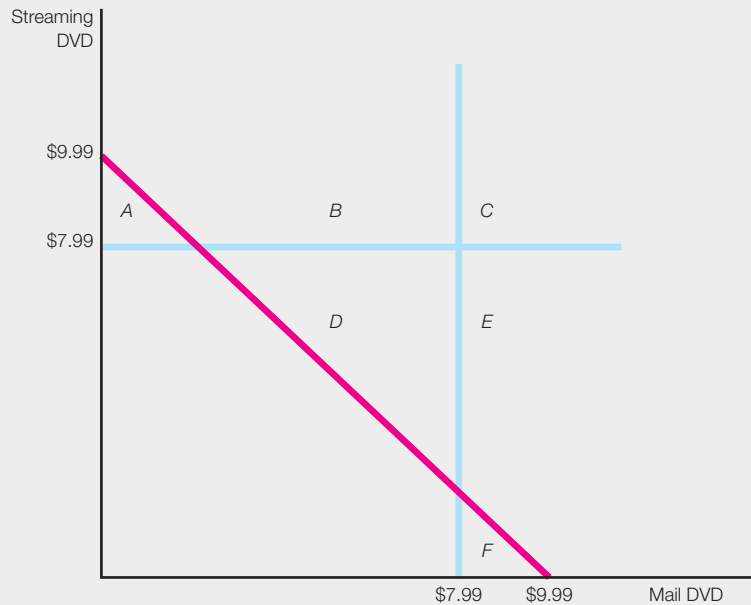
$$\$7.99A + \$7.99B + \$15.98C + \$7.99E + \$7.99F \\ > \$9.99B + \$9.99C + \$9.99D + \$9.99E$$

$$\text{Or } (Z) \quad \$7.99A + \$5.99C + \$7.99F \\ > \$2B + \$9.99D + \$2E$$

The marginal cost of serving the streaming DVDs is very low. The cost of serving the mail DVDs is higher because of postage and handling. The *A* demanders never consume mail DVDs, so their impact in (Z) is strictly on the profit side (assuming the marginal cost of serving them is approximately zero). The *F* customers are new to Netflix, so their net contribution is positive (assuming the \$7.99 they pay exceeds the mailing and handling cost). The costs of handling the *C* customers is the same as before, so the *C* above represents a profit gain. Thus, the left-hand side of (Z) must only be reduced by the cost of serving the new *F* customers to change it from a revenue stream to a profit stream. The *B*, *D*, and *E* customers all previously consumed the bundle. Some probably only streamed (most likely the *B*s) and some probably only mailed (most likely the *E*s). To change the left-hand side of (Z) from a revenue stream to a profit stream, we'd have to subtract out the costs of serving the *B*s, *D*s, and *E*s.

3. If goods' reservation prices are correlated *perfectly* negatively and the marginal cost of production of the goods is zero, pure bundling is best.

4. If goods' reservation prices are negatively correlated, as the marginal cost of production increases, mixed bundling is likely to be better than pure bundling; and as it increases further, separate pricing is likely to be better.



Thus, whether this was a good move or not depends on the number of demanders in each area and the cost of serving them. This is what Netflix managers should have considered in reaching their decision.

Whether they did or not we do not know. We do know managers were forced to publicly reverse the unbundling strategy three months later in October due to a serious drop in demand. Most customers saw the unbundling as a 60% price hike. Even though

Reed Hastings, the CEO of Netflix publicly declared the price hike was necessary, the significant drop in demand forced him to rescind it.

Sources: Tiernan Ray, "Netflix Splits Streaming, DVD Plans, \$15.98 for Both," *Tech Trader Daily*, July 12, 2011, at <http://blogs.barrons.com/techtraderdaily/2011/07/12/netflix-splits-streaming-dvd-plans-15-98-for-both/>; Troy Wolverton, "Netflix Reverses Decision to Split Its Video Streaming and DVD Services," *The Mercury News*, October 11, 2011.

But everything really depends on the reservation prices and the costs of production, so remember that intelligent experimentation is the way to solve bundling pricing. The other approach, especially when the demanders or goods are many, is to use a computer program to search all three types of pricing and every price in each type for the profit-maximizing result.

TYING AT IBM, XEROX, AND MICROSOFT

Tying A pricing technique in which managers sell a product that needs a complementary product.

One form of bundling involving complementary products is called **tying**. Tying is a pricing technique in which managers sell a product that needs a complementary product. The consumer is required, generally by contract, to buy the complementary product from the firm selling the main product. For example, an individual can only download a song from iTunes onto an Apple product. Both the United States and the European Union charged that Microsoft uses a tying strategy to force consumers to use its browser product (Internet Explorer) instead of a rival product (Firefox). Microsoft did this by tying the browser to the Windows operating system, then using its market power to force PC manufacturers to package only Internet Explorer on their machines.

Successful implementation of a tying strategy generally requires the exercise of market power. For example, managers of Apple and Microsoft largely control their markets via their high market shares. Managers engage in tying practices for several reasons. First, it is a way of practicing price discrimination. By setting the price of the complementary product well above its cost, managers can get, in effect, a much higher price from those who use it more often. For example, suppose customer *A* uses a Hewlett-Packard printer to print 10,000 pages per month, whereas customer *B* uses an HP printer to print only 1,000 pages per month. It is hard for Hewlett-Packard to price its machines to obtain more revenue from customer *A*, the more intensive user, than from customer *B*. But if HP can tie the sale of ink cartridges to the sale of its printer, it can get more profit from customer *A* than from customer *B* because it makes more on selling cartridges.

Tying has been challenged as a legal business strategy. Until recently, tying was viewed as per se illegal in most U.S. courts. However, recent decisions suggest judges are recognizing the efficiencies generated by the strategy.

For example, tying can lower transaction costs. Most buyers prefer to purchase a “complete” automobile with all necessary equipment such as tires, engine, and seats. Tying is also used to protect product integrity. Managers are able to control quality levels across product components.

Another reason managers use a tying strategy is to maintain their monopoly position. For example, Microsoft has held a market share of over 90% in the PC operating system market since 1991. Even a competitor the size of IBM was forced to quickly withdraw from this market after spending hundreds of millions of dollars on the OS/2 operating system. Netscape was a concern for Microsoft because its product threatened to reduce the number of application programs written for Windows. So Microsoft wanted to exclude Netscape’s product. When asked the following question by a government lawyer, a Microsoft executive agreed with the lawyer’s assessment:

STRATEGY SESSION: Which is More Expensive: Printers or Ink?

When printers first hit the market, they were expensive. But as more companies produced printers, the price fell. The printer is a one-time expense (for the life of the printer), but the ink cartridges run out of ink and must be replaced. For heavy users, this occurs a lot. So why not entice customers to buy your printer by offering a low price for it, and then after having locked them in, sock it to them by charging high prices for the replacement cartridges, which only work in your printer? This method has been used to sell razors (cheap razor, expensive blades). Well, why wouldn't the customer just refill an empty cartridge with ink? The companies were putting a chip in the cartridges to prevent this, until environmental laws forced recycling of cartridges. Why not use an off-brand cartridge that was engineered for your printer? Because you may be afraid of the off-brand's quality, and you may void the warranty on your printer if you don't use a manufacturer-made cartridge. Hewlett-Packard, with a 40% share of the global market for color printers, had mastered making money from cartridges. One student claims to have bought a new printer when his printer ran out of ink because it was cheaper than buying a new cartridge.

Now Kodak is trying to change all this by introducing an ink-jet printer that costs \$50 more than the competitors' printers but uses ink that is much cheaper than the competitors' ink. Xerox has introduced a color printer that costs \$900 more than the competition but will produce color pages at the same cost as black and white.

Some experts say that the buyers want low-cost hardware (printers) and control costs by controlling the amount of printing done. Kodak feels that people will print more if the cost of ink comes down and reports that its customers are buying nine cartridges per year, while the industry average is four per year. Kodak said that it listened to its customers, who said that high ink costs caused them to limit printing. This may not be good news for trees, but users seem to be happy. A real estate agency that spent \$1 million per year on color printing now reports printing expenses of \$200,000.

Source: Claudia H. Deutsch, "In a Switch, Charging More for Printers and Less for Ink," *New York Times*, September 24, 2007, at <http://www.nytimes.com/2007/09/24/technology>.

And all I am trying to establish is that the reason for that [packaging Internet browser with Windows] was because you believed that if the customer had a choice of the two browsers side by side, the user would, in the vast majority of cases or in the majority of cases, pick Netscape for the reasons that you've identified, correct?⁹

This belief was reiterated in an internal Microsoft e-mail presented at the trial:

It seems clear that it will be very hard to increase browser share on the merits of Internet Explorer alone. It will be more important to leverage

9. From September 21, 1999, trial transcript.

the OS asset to make people—to make people choose Explorer instead of Navigator.¹⁰

In addition, managers may use a tying strategy to ensure the firm's product works properly and its brand name is protected. To do so, the firm insists that customers use its complementary product. For example, Jerrold Electronics Corporation, which installed community antenna systems, required customers to accept five-year maintenance contracts to avoid breakdowns resulting from improper servicing of its equipment. And McDonald's franchises must buy their materials and food from McDonald's so that the hamburgers are uniform and the company's brand name is not tarnished.

TRANSFER PRICING

In some cases, transactions occur where markets do not exist; many times they involve intrafirm pricing. Say there are two divisions of a firm, where a product required as an input is produced exclusively in an upstream plant for use in a product of a downstream plant. Transfer pricing results from creating an internal market that simulates an external one and allows optimal profit-maximizing decisions by managers in both divisions of the firm. For instance, auto companies purchase inputs from their components' divisions so they can produce automobiles. When an external market exists for the product of the upstream division, the rules for the optimal transfer price determination differ because now the upstream division has the option of selling the product in the external market and the downstream division has the option of purchasing the upstream product in the external market.

Transfer pricing is prevalent. A recent survey shows that managers in 91% of the *Fortune* 150 practice transfer pricing and that in one-third of the firms, managers engage in four or more instances per year of intrafirm transactions.

Consider a multidivisional firm with a downstream monopoly and an upstream provider of a component to the downstream product, such as an engine maker serving a downstream automaker. We assume initially there is no external market for engines; that is, no other engine maker can supply engines to the downstream automaker, nor can any other automaker use the engines of the upstream engine maker. Therefore, managers must decide how many engines and autos to make (these are the same because there is no external market for engines).¹¹ The downstream operation is subject to the discipline of the market because autos are sold in an external market. But if there is no external market for the upstream product, what price should change hands between the two divisions to pay for the upstream product? This payment, called a **transfer price**, simulates a market where no formal market exists.

Transfer price Payment that simulates a market where no formal market exists.

10. Ibid.

11. For simplicity, we assume that all the upstream product produced during the period must be sold then. In other words, no inventories of the upstream product can be carried over.

In considering transfer pricing policies, managers need to ensure the profit-maximizing (from the point of view of the entire firm) output of the downstream and the upstream output is produced. Then they must ensure the upstream managers have the right incentive to produce the profit-maximizing amount of the upstream product in the most efficient way.

The following notation enables us to view the transfer pricing issues facing managers. The demand curve for the downstream product is

$$P_D = P_D(Q_D)$$

where P_D is the price of the downstream product per unit, Q_D is in units of the downstream product, and the parentheses mean “function of” in this and the next equation. Recall that the impacts of complementary and substitute goods on this demand are subsumed in the intercept of the demand curve, as shown in Chapter 2.

The production function¹² of the downstream operation is defined as

$$Q_D = f(L_D, K_D | Q_U)$$

The production function is like those of Chapter 5, but it is conditional; that is, it states that Q_D can be produced with labor (L_D) and capital (K_D) *given* the critical upstream input Q_U .

This production function yields a downstream cost function of

$$TC_D = TC_D(Q_D | Q_U)$$

which is the total cost of the downstream division *exclusive* of the cost of the upstream operation.

Finally, the total cost of the upstream division is just a function of Q_U ; it reads

$$TC_U = TC_U(Q_U)$$

and is typical of the cost functions we developed in Chapter 6.

The profit of the multidivisional firm is

$$\pi = TR_D - TC_D - TC_U \quad (10.1)$$

To maximize profit, we must have $\Delta\pi/\Delta Q_U = 0$. Note that Q_U is the variable that controls what managers do. Without the critical input produced by the upstream division, nothing can be produced in the downstream division. And whatever is produced upstream equals the amount produced downstream; that is, $Q_D = Q_U$ when transfer pricing is done correctly. Although we might be tempted to put a total revenue for the upstream division in equation (10.1) ($TR_U = P_U Q_U$, where P_U would be the transfer price), it would be exactly offset by a cost item for the downstream firm (recall that TC_D is the downstream cost *exclusive* of the cost of the upstream product). Because this nets out to zero, it is not included in equation (10.1).

12. The | is not a division sign in this equation or the next one; rather, it reminds us that the functions are conditional on the amount of Q_U .

If we make the left side of equation (10.1) $\Delta\pi/\Delta Q_U = 0$, we must make the following adjustments to the right side of equation (10.1)

$$\frac{\Delta\pi}{\Delta Q_U} = \left(\frac{\Delta TR_D}{\Delta Q_D}\right)\left(\frac{\Delta Q_D}{\Delta Q_U}\right) - \left(\frac{\Delta TC_D}{\Delta Q_D}\right)\left(\frac{\Delta Q_D}{\Delta Q_U}\right) - \frac{\Delta TC_U}{\Delta Q_U} = 0$$

or

$$\left[\left(\frac{\Delta TR_D}{\Delta Q_D}\right) - \left(\frac{\Delta TC_D}{\Delta Q_D}\right)\right]\left(\frac{\Delta Q_D}{\Delta Q_U}\right) = \left(\frac{\Delta TC_U}{\Delta Q_U}\right) \quad (10.2)$$

or

$$(MR_D - MC_D)MP_U = MC_U \quad (10.3)$$

QUANT OPTION

Setting $\partial\pi/\partial Q_U = 0$ will maximize the firm's profit. Thus,

$$\partial\pi/\partial Q_U = [\partial TR_D/\partial Q_D][\partial Q_D/\partial Q_U] - [\partial TC_D/\partial Q_D][\partial Q_D/\partial Q_U] - dTC_U/dQ_U = 0$$

or

$$[[\partial TR_D/\partial Q_D] - [\partial TC_D/\partial Q_D]][\partial Q_D/\partial Q_U] = [dTC_U/dQ_U]$$

or

$$(MR_D - MC_D)MP_U = MC_U$$

Note that $[\partial Q_D/\partial Q_U]$ is the marginal product of the upstream product.

Note that $(\Delta Q_D/\Delta Q_U)$ is the marginal product of the upstream product in producing the downstream product.

The intuition of equation (10.3) is straightforward. If managers produce another unit in the upstream operation, they incur an additional cost, MC_U . Producing that additional upstream product enables the conglomerate to produce MP_U more downstream units. Each additional downstream unit produced causes managers to incur additional cost in the downstream plant (MC_D) but also enables them to earn additional revenue (MR_D). If the additional net revenue earned, $(MR_D - MC_D)MP_U$, which is produced as a result of incurring the additional cost upstream, MC_U , exceeds that additional upstream cost, then managers want to produce the additional unit upstream (because profit increases). If it does not, managers do not want to produce the additional unit upstream (because profit decreases). Managers maximize profit when the additional net revenue earned

downstream as a result of producing an additional unit upstream just equals the additional cost incurred in producing that unit upstream.

But MP_U equals 1 because every time one more unit is produced upstream, one more unit can be produced downstream. In situations where it would appear that the upstream firm has to produce multiple units to enable one additional unit to be produced downstream, such as four tires being required to produce one car, we treat this by requiring one bundle (of four tires) to be produced upstream in order to produce one car. Obviously the situation can also go the other way—such as one extra steer on an upstream cattle ranch enabling the downstream meat processing division to produce $X(>1)$ steaks. Here producing one more steer lets the conglomerate firm produce one more bundle of (X) steaks.

In the situation when $MP_U = 1$, equation (10.3) becomes

$$MR_D - MC_D = MC_U$$

or

$$MR_D = MC_D + MC_U \quad (10.4)$$

and the rule becomes our familiar one: The marginal revenue of the product must equal the marginal cost of producing it. That is, the marginal cost of producing the downstream product is the marginal cost of the downstream operation (remember that this excludes the cost of the upstream operation) plus the marginal cost of the upstream product.

Solving equation (10.4) for $Q_U^* = Q_D^* = Q^*$ gives the correct amount of the upstream product and downstream product produced.

Now suppose managers set the transfer price the downstream division pays and the upstream division receives for its upstream product. They tell the upstream division chief that she will receive P_U for every unit she produces. A profit-maximizing division chief (who is now a price taker) maximizes profit by setting $P_U = MC_U$, as was shown in Chapter 7. But what P_U should managers choose? Clearly it is the P_U that results in Q_U^* units being produced. This is shown in Figure 10.9.

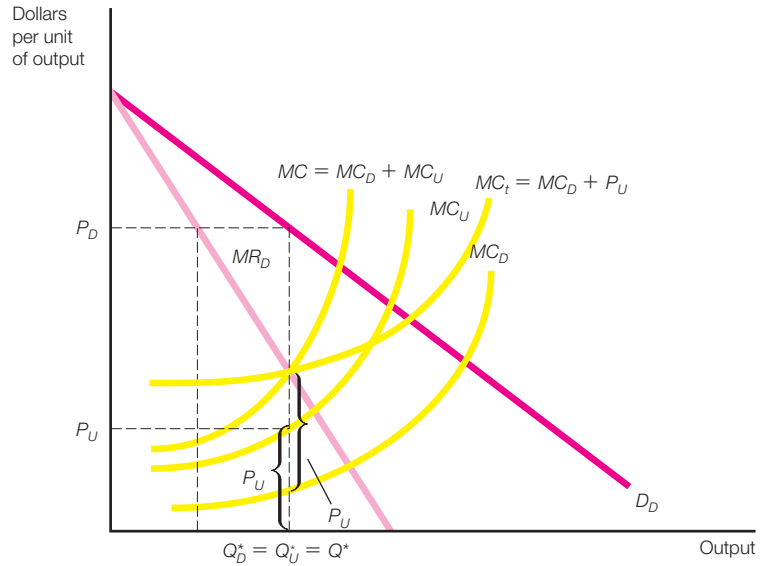
What difference does it make what P_U managers choose? Whatever it is, would it not merely cause the upstream division to have $P_U Q_U^*$ in revenues and the downstream division to have $P_U Q_U^*$ in costs? The two terms merely cancel each other out (which is why we left them out of equation (10.1)). From the point of view of the conglomerate, profit is the same. However, *the profit of each division differs*. And because managers' bonuses are often predicated on their division's profit, these managers care about that transfer price. If the conglomerate's managers determine the optimal Q (that is, Q^*) and order both divisions to produce it, then the conglomerate maximizes profit regardless of the transfer price.

But if the managers are trying to maximize firm profit, it is critical that the correct P_U be chosen. If P_U is set too high, the upstream division will produce

FIGURE 10.9

Determination of the Transfer Price, Given No External Market for the Transferred Good

The optimal transfer price, P_U , equals the marginal cost at the optimal output, Q^* .



too much of the product (recall that in price-taking situations, the marginal cost is rising). Also, managers of the downstream division will see its marginal cost of producing another unit ($MC_D + P_U = MC_t$) as too high and therefore will produce too little downstream output. Therefore, profit will not be maximized. If P_U is set too low, managers of the downstream division will see their marginal cost of producing another unit as too low and will therefore want to produce more than the optimal output (which they cannot do because the upstream division has produced less of its output).

TRANSFER PRICING: A PERFECTLY COMPETITIVE MARKET FOR THE UPSTREAM PRODUCT

In many cases there is a market outside the firm for the product transferred from one division to the other. If this is true, the output levels of the downstream and upstream divisions no longer need be equal. If the downstream division wants more of the upstream product than is produced by the upstream division, it can buy some from external suppliers. If the upstream division produces more of its

product than the downstream division wants, it can sell some to external customers. Assuming the market for the upstream product is perfectly competitive, we can readily determine how managers should set the transfer price.

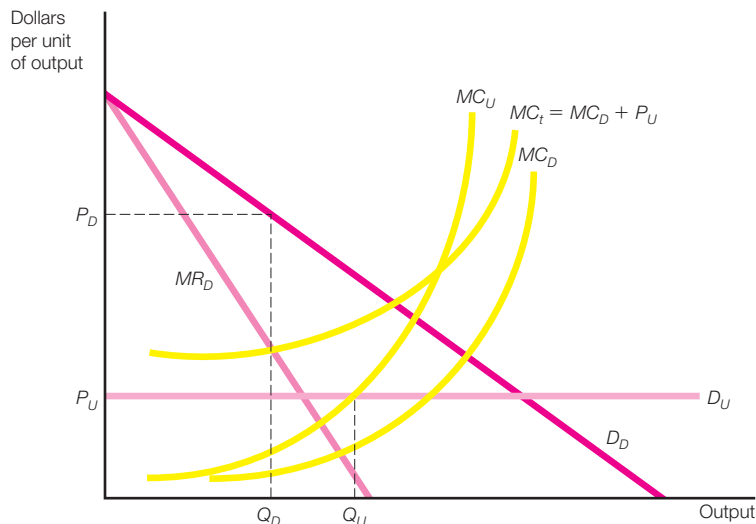
Figure 10.10 shows the optimal price and output for the firm as a whole. Because there is a perfectly competitive market for the upstream product, managers at the upstream division act as if they see a horizontal demand curve, D_U , where the price is P_U , the price of the upstream product in the external market. To maximize profit, managers at the upstream division should produce the output Q_U , where the marginal cost of the upstream division, MC_U , equals the *externally* determined market price P_U . In this sense the upstream division behaves like a perfectly competitive firm.

To maximize the firm's overall profit, the transfer price should equal P_U , the price of the upstream division in the perfectly competitive market outside the firm. Because managers at the upstream division can sell as much product as they want to external customers at a price of P_U , they have no incentive to sell it at a price below P_U to the downstream division. Similarly, because managers at the downstream division can buy as much of the upstream product as they want from external suppliers at a price of P_U , they have no incentive to buy it from the upstream division at a price above P_U .

FIGURE 10.10

Determination of the Transfer Price, Given a Perfectly Competitive External Market for the Transferred Product

The optimal transfer price, P_U , equals the market price of the transferred product.



STRATEGY SESSION: Transfer Pricing in Practice

Transfer pricing plays two different roles for managers. One is the proper allocation of resources within a conglomerate business when goods or services are transferred from one division of the conglomerate to another. Since we view prices as a signal to managers as part of the process of how much to produce and how to produce it, we should want those signals to the division managers to be such that the overall profit of the conglomerate is maximized.

With that in mind, Michalski, Bunger, and Stiller investigated the role of transfer prices in a cradle-to-grave analysis from producing hydrogen for fueling automobiles to the retail outlet that would dispense the fuel to automobiles. Unless the process was treated as one system, suboptimal transfer prices at the various handoffs of product through the production and distribution phases would impact the profitability of the whole process (and the profitability of any constituent part). This is not surprising given the analysis in the text.

The other role is in tax policy. On January 20, 2011, Ernst & Young issued its *2010 Global Transfer Pricing Survey: Addressing the Challenges of Globalization*. The main finding from surveying 877 multinational enterprises in 25 countries was that transfer pricing remains the number one challenge for these leading companies. Governments, facing record deficits, are stepping up their scrutiny and enforcement of transfer price practices or introducing transfer pricing rules for the first time. KPMG conducts its own annual survey (*Global Transfer Pricing Review*), and notes that increasing international trade is putting pressure on the industry to standardize transfer pricing rules. The KPMG report is over 200 pages long, documents the policies in 64 countries, and gives snapshots of developments in 35 countries that either have no policies or are just beginning to develop their policies.

Just how well have governments been in stepping up enforcement and stopping the evasion of

taxes via the use of transfer prices? A National Bureau of Economic Research paper^a looks at the use of transfer prices as well as other mechanisms, such as using hybrid entities that are treated as corporations in some countries but flow-through entities in others, to transfer profits from high-tax countries to low-tax countries. The paper finds that in the last 20 years, multinational corporations have taken to “domicile swapping” with a passion. For nine developed countries with sufficient observations over the period from 1989 to 2009, the effective tax rate (ETR) had a mean (and median) decline of 12 percentage points. In Japan, the decline was 22 percentage points; in Switzerland and Great Britain, 15; and in the United States, 12. Bloomberg estimates that Google (domiciled in Bermuda) saved \$3.1 billion in the last three years by such maneuvering.

Despite all of this opportunity (ETRs in high-tax countries are double the ETRs in low-tax countries), some companies stay domiciled in high-tax countries. Japanese-domiciled companies face the highest effective tax rate, followed by the United States, France, and Germany. Although the ETR has declined as mentioned above, the rank ordering of countries by ETRs has changed little over time.

^a Kevin S. Markle and Douglas A. Shackelford, “Cross-Country Comparisons of Corporate Income Taxes,” Working Paper No. 16839.

Source: Jan Michalski, Ulrich Bunger, and Christoph Stiller, “Business Analysis of the Hydrogen Refueling Station Infrastructure and the Role of the Transfer Pricing System,” *International Journal of Hydrogen Energy* 36 (2011), pp. 8152–8157; Terry Hayes, “Transfer Pricing Remains Key Challenge for MNEs,” *Journal of International Taxation* 22, no. 4 (April 2011), p. 8; KPMG, “Global Transfer Pricing Review” at www.snipurl.com/27pakr; Mythili Bhusnurmath, *Economic Times of India* (New Delhi), “MNCs Lower Tax Burden by Swapping Domicile,” May 15, 2011, at <http://economictimes.indiatimes.com/opinion/policy/mcns-lower-tax-burden-by-swapping-domicile/articleshow/8309924.cms>.

Managers at the downstream division, which must buy the upstream product at price P_U regardless of where it comes from, have a marginal cost of MC_t , which is the sum of the downstream division's marginal cost, MC_D , and the market-determined price of the upstream product, P_U . To maximize their own profit, managers at the downstream division must choose the output level, Q_D , where their marginal cost, $MC_t (= MC_D + P_U)$, equals their marginal revenue, MR_D . Figure 10.10 shows that the output of the downstream division, Q_D , is less than the output of the upstream division, Q_U ; so the optimal solution in this case calls for the conglomerate's upstream division to sell part of its output (specifically $Q_U - Q_D$ units) to outside customers.¹³

THE GLOBAL USE OF TRANSFER PRICING

Transfer pricing is widespread. Many firms have policies whereby one division can buy another division's product, with the transfer price determined by various means. This observation found support in a 1992 survey of transfer pricing that targeted *Fortune* 500 firms. For domestic interdivisional transfers, the most common methods were the use of market prices, actual or standard full production costs, full production costs plus a markup, and negotiated prices. For international transfers, market-based transfer prices and full production costs plus a markup were the most commonly reported methods. Comparing the results to an earlier survey conducted in 1977, the shift has been to market-based prices in both the domestic and international markets.¹⁴

Managers use transfer pricing to shift profits between divisions to minimize tax liability. This, done on a state-by-state and country-by-country basis, has caused government officials to investigate transfer pricing as a method of avoiding taxation. A 1999 survey by Ernst & Young showed that the number one international tax issue is transfer pricing. Firms are concerned with double taxation and onerous penalties for noncompliance. Many countries have enacted legislation enabling their tax agencies to intensify their transfer pricing inquiries and regulation enforcement. These countries feel that managers use transfer prices to decrease profit in high-tax countries, transferring this profit to low-tax countries. Items included in transfer pricing are goods, services, property, loans, and leases. Fortunately (and in line with the theory developed in this chapter), survey respondents noted that "the most important factor shaping transfer pricing policies is maximization of operating performance, not optimizing tax arrangements."¹⁵

Suppose the tax rate in a downstream country is α and the tax rate in an upstream country is β where $\alpha > \beta$. Suppose the case is one of no external market for the upstream product. The after-tax profit in the downstream country is

$$(1 - \alpha)(TR_D - TC_D - P_U Q_U)$$

13. Of course it is not always true that Q_D is less than Q_U . Whether this is the case depends on the shape and position of the marginal cost curves (MC_D and MC_U) and the demand curve as well as the price of the transferred product in the external perfectly competitive market. If $Q_D > Q_U$, the downstream division purchases the required $Q_D - Q_U$ units in the external market at the market-determined price of P_U .

14. Roger Tang, "Transfer Pricing in the 1990s," *Management Accounting* vol. 73(8), pp. 22–26. This 1992 survey replicated one conducted in 1977.

PROBLEM SOLVED: The Orion Corporation

Consider the Orion Corporation, where an upstream chemical division [P] produces a product that it transfers to a downstream marketing division [M] which packages the basic chemical into the final product and sells it to outside customers. To illustrate how managers should calculate the optimal output rates, assume demand and cost conditions are as follows. The demand for the finished product sold by Orion's downstream marketing division is

$$P_M = 100 - Q_M \quad (10.5)$$

where P_M is the price (in dollars per ton) of the finished product and Q_M is the quantity demanded (in millions of tons per year). Excluding the cost of the basic chemical, the marketing division's total cost function is

$$TC_M = 200 + 10Q_M \quad (10.6)$$

where TC_M is the division's total cost (in millions of dollars).

Turning to Orion's upstream production division, its total cost function is

$$TC_P = 10 + 2Q_P + 0.5Q_P^2 \quad (10.7)$$

where TC_P is total production cost (in millions of dollars) and Q_P is the total quantity produced of the basic chemical (in millions of tons per year). As we did earlier, we assume a perfectly competitive market for the basic chemical (the upstream output). Assume that its price in this market is \$42 per ton.

Under these conditions, managers can readily determine the optimal rate for each division as well as the proper transfer price for the basic chemical. The production division can sell all the basic chemical that it wants at \$42 per ton. Therefore, its marginal revenue equals \$42. From equation (10.7), we see that $\Delta TC_P / \Delta Q_P = MC_P = 2 + Q_P$

To find the output that maximizes the production division's profit, managers set its marginal revenue equal to its marginal cost

$$MR_P = 42 = 2 + Q_P = MC_P$$

or

$$Q_P = 40$$

Hence the production division should produce 40 million tons per year of the basic chemical.

QUANT OPTION

Orion's marginal cost is $dTC_P/dQ_P = 2 + Q_P$.

15. "Multinationals Face Greater Transfer Pricing Scrutiny According to New Ernst & Young Survey," *Business Wire*, November 3, 1999.

and the after-tax profit in the upstream country is

$$(1 - \beta)(P_U Q_U - TC_U)$$

The transfer price of the basic chemical should be its price in the perfectly competitive market outside the firm. This market price is \$42 per ton, and the transfer price should be the same. Also, we know from our earlier work that the marketing division's marginal cost, MC_t , is the sum of its own marginal marketing cost, MC_M , and the transfer price. That is,

$$MC_t = MC_M + P_U$$

where $P_U = \$42$ and its own marginal marketing cost equals $MC_M = \Delta TC_M / \Delta Q_M$.

From equation (10.6), we see that $\Delta TC_M / \Delta Q_M = MC_M = 10$. Therefore,

$$MC_t = 10 + 42 = 52$$

To maximize the marketing division's profit, managers must set its marginal cost equal to its marginal revenue. The marketing division's total revenue is

$$TR_M = P_M Q_M = (100 - Q_M) Q_M = 100 Q_M - Q_M^2$$

The marketing division's marginal revenue is therefore

$$\Delta TR_M / \Delta Q_M = 100 - 2Q_M$$

Setting this expression for its marginal revenue equal to its marginal cost, we find that

$$MR_M = 100 - 2Q_M = 52 = MC_t = MC_M + P_U$$

or

$$Q_M = 24$$

Hence the marketing division should sell 24 million tons per year of the base chemical at a price of \$76; that is, $P_M = 100 - 24$.

To sum up, managers at the Orion Corporation's production division should produce 40 million tons per year of the basic chemical. Of this amount, 24 million tons should be transferred to Orion's marketing division at the market price of \$42 per ton, and 16 million tons (40 - 24) should be sold externally at the market price of \$42 per ton. The transfer price should be the same as the market price: \$42 per ton.

QUANT OPTION

Take this at face value.

Orion's marginal cost is $dTC_M / dQ_M = 10$; Orion's marginal revenue is $dTR_M / dQ_M = 100 - 2Q_M$.

Suppose all profits are expressed in the same currency; that is, we have adjusted for exchange rates and the P_U was set to maximize before-tax profit. The overall conglomerate's after-tax profit is

$$\begin{aligned}
& (1 - \alpha)(TR_D - TC_D) - (1 - \beta)(TC_U) + (1 - \beta - [1 - \alpha])(P_U Q_U) \\
& = (1 - \alpha)(TR_D - TC_D) - (1 - \beta)(TC_U) + (\alpha - \beta)(P_U Q_U)
\end{aligned}$$

Because $\alpha > \beta$, the conglomerate's after-tax profit is higher if P_U is greater. But the optimal before-tax profit-maximizing P_U is what it is (and it could be low).

Suppose the firm, having determined the optimal $Q_D^* = Q_U^* = Q^*$, now creates a "subterfuge P_U " = P_U^S for tax purposes and sets it such that

$$P_U^S = (TR_D - TC_D)/Q_U^*$$

With this P_U^S , the after-tax profit in the downstream country becomes 0 and the conglomerate's after-tax profit is

$$(1 - \beta)(TR_D - TC_D - TC_U)$$

That is, all corporate profit is taxed at the lowest tax rate. Here we see the motivation of high tax rate countries to look at the transfer price policies of multinational firms.

Just why have transfer prices become so important on the international level? Four basic reasons exist: increased globalization, different levels of taxation in various countries, greater scrutiny by tax authorities, and inconsistent rules and laws in the various tax jurisdictions. Transfer price policies that seem to cause the fewest legal problems in the international scenario are (1) comparable uncontrolled price, in which the prices are the same or similar to "arm's-length" transaction prices; (2) cost-plus prices, in which a markup used in arm's-length transactions is added to the seller's cost of the good or service; and (3) resale price, in which the resale price is used as a base for determining an arm's-length margin for the functions performed by the selling company.¹⁶

SUMMARY

1. Bundling is a strategy that enables managers to increase profit by selling two or more goods in a bundle. In general, bundling works better if the reservation prices of goods are negatively correlated; that is, one group has a high reservation price for one good and a low reservation price for another good relative to another group. Even so, bundling need not generate more profit relative to pricing each good separately. A pure bundle occurs when the goods are sold only as a bundle. Mixed bundling occurs when goods are sold both as a bundle and at least one good is sold separately. No analytical model is available to solve the bundling pricing problem, so experimentation or a computer model is used. Managers also use bundling as an entry deterrent.

2. Many large firms are multidivisional, and one division sells its product to another division of the firm. To maximize the firm's overall profit, it is important that the price at which this transfer takes place, the so-called transfer

16. Brenda Humphreys, "International Transfer Pricing: More Important Than Ever Before!" *Cost & Management* vol. 68(4), pp. 24–26.

price, be set properly. If there is no market outside the firm for the transferred product, the transfer price should equal the marginal production cost of the transferred product at the optimal output. If the outside market is perfectly competitive, the transfer price should equal the market price. Transfer prices have been used by global firms to transfer profits from high-tax countries to low-tax countries.

PROBLEMS

wwwnorton.com/studyspace 

1. The Locust Corporation is composed of a marketing division and a production division. The marginal cost of producing a unit of the firm's product is \$10 per unit, and the marginal cost of marketing it is \$4 per unit. The demand curve for the firm's product is

$$P = 100 - 0.01Q$$

where P is the price per unit (in dollars) and Q is output (in units). There is no external market for the good made by the production division.

- a. How should managers set the optimal output?
 - b. What price should managers charge?
 - c. How much should the production division manager charge his counterpart in marketing for each unit of the product?
2. The Xerxes Company is composed of a marketing division and a production division. The marketing division packages and distributes a plastic item made by the production division. The demand curve for the finished product sold by the marketing division is

$$P_0 = 200 - 3Q_0$$

where P_0 is the price (in dollars per pound) of the finished product and Q_0 is the quantity sold (in thousands of pounds). Excluding the production cost of the basic plastic item, the marketing division's total cost function is

$$TC_0 = 100 + 15Q_0$$

where TC_0 is the marketing division's total cost (in thousands of dollars). The production division's total cost function is

$$TC_1 = 5 + 3Q_1 + 0.4Q_1^2$$

where TC_1 is total production cost (in thousands of dollars) and Q_1 is the total quantity produced of the basic plastic item (in thousands of pounds). There is a perfectly competitive market for the basic plastic item, the price being \$20 per pound.

- a. What is the optimal output for the production division?
- b. What is the optimal output for the marketing division?

- c. What is the optimal transfer price for the basic plastic item?
 - d. At what price should the marketing division sell its product?
3. Knox Chemical Corporation is one of the largest producers of isopropyl alcohol, or isopropanol, as it frequently is called. Isopropanol is used to produce acetone, an important industrial chemical; it is also used to make various chemical intermediate products. Because Knox Chemical produces both acetone and these chemical intermediates, it uses much of the isopropanol it makes. One of the many tasks of Knox's product manager for isopropanol is to set transfer prices for isopropanol within the company.
- a. Knox's product manager for isopropanol generally sets the transfer price equal to the prevailing market price. Is this a sensible procedure?
 - b. When the production of phenol expands rapidly, a great deal of acetone is produced because it is a by-product of the process leading to phenol. What effect do you think this has on the market price of isopropanol?
 - c. In producing a pound of phenol, 0.6 pound of acetone is produced. Are phenol and acetone joint products?
 - d. Are they produced in fixed proportions?¹⁷
4. The reservation prices (in dollars) of three classes of demanders (classes A, B, and C) for Ricky Parton's (a Latin country-western singer) compact discs are given in the table that follows:

Class	CD 1	CD 2
A	11	5
B	8	9
C	9	10

It costs \$4 to produce and distribute each compact disc. The company can sell each CD separately, can put them together as a boxed set (that is, as a pure bundle), or can sell them in a mixed bundling format (offer the CDs both separately and as a boxed set). Assume that each demander wants only one of each of the CDs at the reservation price (or at any lower price) and that there are an equal number of demanders in each class. For simplicity, assume that the only costs are those mentioned here.

- a. What pricing method would you advise Ricky's company to use?
 - b. How much better (profitwise) is the best pricing method than the second most profitable pricing method?
5. Bob and Ron's Stereo sells televisions and DVD players. They have estimated the demand for these items and have determined that there are three consumer types (A, B, and C) of equal number (assume one for simplicity) that have the following reservation prices for the two products. Bob and Ron's cost

17. For further discussion, see E. R. Corey, *Industrial Marketing: Cases and Concepts*, 3d ed. (Englewood Cliffs, NJ: Prentice-Hall, 1983).

for a TV is 9 and for a DVD player is 9. It will cost Bob and Ron 18 to produce a bundle of one TV and one DVD player.

Consumer	TV	DVD Player
A	28	12
B	29	4
C	30	10

Any consumer's reservation price for a bundle of one TV and one DVD player is the sum of their reservation prices for each item. Consumers will demand (at most) one TV and one DVD player.

- a. If Bob and Ron only consider pricing each item separately, pricing a pure bundle, or pricing a mixed bundle as their pricing policy, what price(s) would maximize their profit and what would be their profit?
 - b. If Bob and Ron were able to perfectly price discriminate (that is charge different prices to different consumers, how much would their profit increase over their optimal profit in part a?)
6. The University of Pennsylvania basketball team will play both the University of Kansas and Nowhere University this year on Penn's campus. Kansas is a nationally ranked team, while Nowhere is just plain terrible.

The athletic director traditionally prices each game separately. You approach him and point out that two other pricing options exist. One possibility is to offer a pure bundle, that is, a ticket package containing one Kansas ticket and one Nowhere ticket. The second possibility is a mixed bundle. In this situation, a pure bundle is offered but admissions to the games can also be sold separately. It costs Penn a constant 5 per spectator to produce a game. It would cost Penn 10 to produce a bundle of a Kansas game and a Nowhere game.

Three types of potential spectators exist (A, B, and C). There are an equal number of types (for simplicity, assume one of each type). Their reservation prices for each game are shown below:

Spectator	Kansas	Nowhere
A	40	13
B	49	3
C	3	30

Penn's policy is not to price discriminate. A spectator's reservation price for a bundle of the two games is the sum of their reservation prices for each game. A spectator wants (at most) one admission to each game.

- a. What's your pricing advice to the athletic director (so that the director maximizes Penn's profit)?
 - b. Given the current pricing policy of Penn, what's your advice worth to the athletic director?
7. GeeM has a sporty wheel package and a luxury interior package that it is considering offering to its auto buyers. GeeM has estimated that there are three consumer types (*A*, *B*, and *C*—all of equal magnitude—for simplicity, consider it one of each type). Consumers want (at most) one of each package. It costs GeeM 5 to produce a sporty wheel package and 10 to produce a luxury interior package. It will cost GeeM 15 to produce a bundle consisting of both packages. The following are the consumer reservation prices for each package:

Consumer	Wheels	Interior
<i>A</i>	11	24
<i>B</i>	35	12
<i>C</i>	18	28

A consumer's reservation price for a bundle consisting of sporty wheels and a luxury interior is the sum of the individual component reservation prices. GeeM does not price discriminate.

GeeM has solicited your help in pricing the wheel and interior package. You know that they could sell the packages separately, as a pure bundle, or as a mixed bundle.

Of those three pricing strategies, which one would maximize GeeM's profit? What are the prices (what is the price) that you suggest? How much better is the best pricing strategy than the second best pricing strategy?

8. Food for Life makes health foods for active, outdoor people. Their three basic products are whey powder, a high protein strength bar, and a meal additive that has the taste and consistency of sawdust. Research shows that consumers fall into two types (*A* and *B*) and these are described in the table below by their reservation prices for the products. Each consumer will demand no more than one unit of any product at their reservation price. The consumers will value a bundle of the products at the sum of the constituent reservation prices. Each product costs \$3 to produce. A bundle of all three products costs \$9 to produce. Food for Life does not price discriminate.

Consumer	Whey	Strength	Sawdust
<i>A</i>	10	16	2
<i>B</i>	3	10	13

There is an equal number of each consumer type (for simplicity, one of each type).

What pricing (profit-maximizing) strategy (among pricing separately, pure bundling, and mixed bundling) would you recommend to Food for Life? Why? Only bundles of all three products need to be considered.

EXCEL EXERCISE: TRANSFER PRICING

Suppose that the demand for the downstream product is given by

$$P_D = 100 - 0.1Q_D$$

The total cost of the downstream division (exclusive of the cost of the upstream input) is

$$TC_D = 5Q_D$$

The total cost of the upstream division is given by

$$TC_U = 20Q_U + 0.05Q_U^2$$

It takes one unit of the upstream input to make one unit of the downstream output; the upstream division cannot sell its product to anyone but the downstream division, and the downstream division can only buy its input from the upstream division. The upstream division's product is unique, and there is no other producer of their product.

The downstream total revenue is

$$TR_D = P_D * Q_D = (100 - 0.1Q_D) * Q_D = 100Q_D - 0.1Q_D^2$$

The downstream marginal revenue is

$$MR_D = dTR_D/dQ_D = 100 - 0.2Q_D$$

The downstream marginal cost (exclusive of the cost of the upstream input) is

$$MC_D = dTC_D/dQ_D = 5$$

The upstream marginal cost is

$$MC_U = dTC_U/dQ_U = 20 + 0.1Q_U$$

To maximize profit, the conglomerate firm (combination of upstream and downstream) will set $MR_D = MC_D + MC_U$, that is,

$$MR_D = 100 - 0.2Q_D = 5 + 20 + 0.1Q_U = 25 + 0.1Q_U = MC_D + MC_U$$

or $100 - 0.2Q = 25 + 0.1Q$ since $Q_D = Q_U = Q$ because it takes one unit of input to make one unit of output.

Thus $0.3Q = 75$

or $Q = 250$

Substituting $Q = Q_D = 250$ into the downstream demand gives

$$P_D = 100 - 0.1 \cdot 250 = 100 - 25 = 75$$

Substituting $Q = Q_D = 250$ into the upstream marginal cost function gives the optimal transfer price that the downstream division should pay the upstream division for a unit of the upstream division's output, that is,

$$P_U = MC_U = 20 + 0.1 \cdot 250 = 20 + 25 = 45$$

The total revenue of the downstream division is

$$TR_D = P_D \cdot Q_D = 75 \cdot 250 = 18,750$$

The total cost of the downstream division (exclusive of the cost of the upstream division) is

$$TC_D = 5 \cdot 250 = 1,250$$

The cost of the downstream division of obtaining the input from the upstream division is

$$P_U \cdot Q_U = 45 \cdot 250 = 11,250$$

The profit of the downstream division is

$$\pi_D = TR_D - TC_D - P_U \cdot Q_U = 18,750 - 1,250 - 11,250 = 6,250$$

The total revenue of the upstream division is

$$TR_U = P_U \cdot Q_U = 45 \cdot 250 = 11,250$$

The total cost of the upstream division is

$$\begin{aligned} TC_U &= 20Q_U + 0.05Q_U^2 = 20 \cdot 250 + 0.05 \cdot 250^2 = 5,000 + 0.05 \cdot 62,500 \\ &= 5,000 + 3,125 = 8,125 \end{aligned}$$

The profit of the upstream division is

$$\pi_U = TR_U - TC_U = 11,250 - 8,125 = 3,125$$

The profit of the combined firm is

$$\pi = \pi_D + \pi_U = 6,250 + 3,125 = 9,375$$

Suppose that the transfer price was set at 65 instead of the optimal 45. What would happen? Then the marginal cost of the input to the downstream plant would be 65 and the downstream would optimize where

$$MR_D = 100 - 0.2Q_D = 5 + 65 = 70 = MC_D + MC_U = MC_D + P_U$$

or

$$0.2Q_D = 30$$

or

$$Q_D = 150$$

Substituting $Q_D = 150$ into the downstream demand gives

$$P_D = 100 - 0.1 \cdot 150 = 100 - 15 = 85$$

The total revenue of the downstream division is

$$TR_D = P_D \cdot Q_D = 85 \cdot 150 = 12,750$$

The total cost of the downstream division (exclusive of the cost of the upstream input) is

$$TC_D = 5 \cdot 150 = 750$$

The cost of the downstream division of obtaining the input from the upstream division is

$$P_U \cdot Q_U = 65 \cdot 150 = 9,750$$

The profit of the downstream division is

$$\pi_D = TR_D - TC_D - P_U \cdot Q_U = 12,750 - 750 - 9,750 = 2,250$$

Managers of the upstream division will only produce 150 units, which is the output demanded by downstream managers. (The output is worthless to managers of external firms.)

The total revenue of the upstream division is

$$TR_U = P_U \cdot Q_U = 65 \cdot 150 = 9,750$$

The total cost of the upstream division is

$$\begin{aligned} TC_U &= 20Q_U + 0.05Q_U^2 = 20 \cdot 150 + 0.05 \cdot 150^2 \\ &= 3,000 + 0.05 \cdot 22,500 = 3,000 + 1,125 = 4,125 \end{aligned}$$

The profit of the upstream division is

$$\pi_U = TR_U - TC_U = 9,750 - 4,125 = 5,625$$

The profit of the combined firm is

$$\pi = \pi_D + \pi_U = 2,250 + 5,625 = 7,875$$

By setting the wrong transfer price the combined companies' profits fell by $9,375 - 7,875 = 1,500$.

Here's how we get the same results with a spreadsheet. Suppose the following data was available about the market price and quantity demanded and the upstream costs associated with quantity produced.

These numbers can be derived from the demand curve for the downstream product ($P_D = 100 - 0.1Q_D$) and the total cost curve for the upstream product ($TC_U = 20Q_U + 0.05Q_U^2$). Managers often use such discrete models in estimating demand.

Price	Downstream Quantity Demanded	Upstream Total Cost
100	0	0
95	50	1,125
90	100	2,500
85	150	4,125
80	200	6,000
75	250	8,125
70	300	10,500
65	350	13,125
60	400	16,000
55	450	19,125
50	500	22,500

You know that the marginal cost of the downstream equals its average variable cost equals 5.

Call up your spreadsheet. Enter 100 in cell A1, 95 in cell A2, and so on, until you enter 50 in cell A11.

Enter 0 in cell B1, 50 in cell B2, and so on until you enter 500 in cell B11. Columns A and B give the downstream demand curve.

Enter $=A1*B1$ in cell C1. Then click on the lower right hand corner of cell C1 and drag your mouse until you reach cell C11. Column C will be the total revenue of the downstream division.

Enter $=5*B1$ in cell D1. Then click and drag until cell D11. Column D is the downstream division's total cost (exclusive of the cost of the upstream division).

Enter 0 in cell E1, 1125 in cell E2, and so on, until you enter 22,500 in cell E11. Column E is the upstream total cost.

Enter $=D1+E1$ in cell F1. Then click and drag until you reach cell F11. Column F is the combined firm's total cost.

Enter $=C1-F1$ in cell G1. Then click and drag until you reach cell G11. Column G is the combined firm's profit. Search column G for the highest number (or enter $=\text{Max}(G1:G11)$ in cell G12 and let the spreadsheet find it). The highest combined profit will be 9,375 in cell G6. Reading leftward on the sixth row, you will see that this entails producing 250 units (cell B6) and pricing the downstream product at 75 (cell A6).

Enter $=(E2-E1)/(B2-B1)$ in cell H2. Then click and drag until you reach cell H11. The $(E2-E1)$ type of calculation gives the firm's discrete marginal cost for 50 increment increases in output. Dividing $E2-E1$ by $B2-B1$ gives an estimate of the per-unit marginal cost in that increment. At the optimal output, the marginal cost for producing the 50 units between 200 and 250 is 2,125. Dividing 2,125

by 50 gives 42.5 as an average marginal cost of producing a unit. The marginal cost of producing the 50 units between 250 and 300 is 2,375. Dividing 2,375 by 50 gives 47.5 as an average marginal cost of producing a unit. Thus, if the upstream transfer price for a unit of upstream product had been set at 47.5, too many units would have been produced (300). But with the transfer price set at 42.5, the optimal 250 units would be produced. Why is our optimal transfer price 42.5, different from the 45 price derived from the above calculus model? This model is discrete in nature; instead of looking at small changes, our change in demand is in 50 unit intervals. Though, note both models identify optimal upstream and downstream output of 250.

Suppose we had more granular data and knew the upstream cost of producing 249 units was 8,080.05 (as shown in the upstream cost formula at the beginning of the problem). Then the discrete marginal cost of producing the 250th upstream product would be $8,125 - 8,080.05 = 44.95$, which is virtually the 45 shown in the earlier calculus case.

PART 6

THE STRATEGIC WORLD OF MANAGERS

CHAPTER 11

LEARNING OBJECTIVES

- Cooperative Behavior
- The Breakdown of Collusive Agreements
- Price Leadership
- Possible Behavior in Markets with Few Rivals
- Duopolists and Price Competition with Differentiated Products
- The Sticky Pricing of Managers
- Summary
- Problems
- Excel Exercise: Dominant Firm Price Leader
- Excel Exercise: Cournot
- Excel Exercise: Stackelberg

Oligopoly A market with a small number of firms.

OLIGOPOLY

Now we come to our last market structure, that of oligopoly. An **oligopoly** is a market with a small number of firms. As a general rule, you'd like to manage as an oligopolist; they realize relatively high profits. Think of the U.S. petroleum industry, where a few firms account for most of the industry's refining capacity. Oligopolies are strategically interesting from the managerial view. There is a tight interdependence between managers of rival firms because of the small set of players. This causes managers to explicitly consider the reactions of rivals in formulating pricing policy. When managers at Exxon Mobil raise their price of home heating oil by \$0.01 or \$0.02 per gallon, they try to anticipate the reaction of rival managers. If rivals decide against such a price increase, it is likely that Exxon Mobil managers will rescind the cut; otherwise, those rivals will capture a significant number of Exxon Mobil customers. In the next chapter, we offer managers game theory as a guide to this process of making strategy.

Oligopolies are global phenomena. For example, the market for commercial aircraft is dominated by Boeing and Airbus. Victoria Thieberger of *The Financial Times* writes about a duopoly in the Australian grocery market. She notes that two companies, Woolworth Ltd. and Wesfarmers Ltd., control over 80% of the grocery sector.¹

There are many reasons why oligopolists are able to rule markets for scores of years. One is a high entry barrier that managers erect using their cooperative market power. Managers at smaller competitors claimed the grocery duopolists were

1. Victoria Thieberger, "Costco Plans Australia Foray to Challenge Duopoly," at www.reuters.com, June 24, 2008.

using their market power to negotiate contracts that discouraged landlords from renting space to them. Thieberger also notes that managers at both companies were part of an inquiry into the setting of eerily similar prices at the rival stores.

Government fiat is another reason for duopolies. The U.S. petroleum industry was once a monopoly controlled by John D. Rockefeller. A brilliant strategist, Rockefeller transformed the industry into a design of his own making. The courts finally tore asunder what rivals could not breach.

A more common reason for oligopolies is economies of scale. Because costs decrease as output expands, only a few firms can survive in the market. Managers of these firms still achieve cost savings even when their output represents a substantial percentage of the market. Scale economies were a large part of Rockefeller's success. Even the courts could not use their power to significantly fragment the industry because of the underlying economics.

The hallmark of oligopoly strategy is its behavioral nature. In contrast to perfect competition or monopoly, for which there is a single unified model, behavior is more varied in oligopolies. This variance in behavior is due to the tight interdependence between market rivals.

COOPERATIVE BEHAVIOR

Conditions in oligopolistic industries tend to encourage cooperation among rival managers. This can increase profit, decrease uncertainty, and raise barriers to discourage others from entering the market. However, maintaining cooperative behavior is difficult. There are usually incentives for cooperative parties to "cheat"; and in most countries formal collusive agreements are not enforceable.

If a collusive arrangement is made openly and formally, it is called a **cartel**. In some countries cartels are legally acceptable; but in the United States most collusive agreements, whether secret or open cartels, were outlawed by the Sherman Antitrust Act (discussed in detail in Chapter 17), which dates back to 1890. But this does not mean the government does not see the cooperative value of oligopolies. Major League Baseball is exempted from the U.S. antitrust laws by an act of Congress. And, governments have allowed international air carriers to form alliances to set prices so managers can use economies of scale to reduce costs.

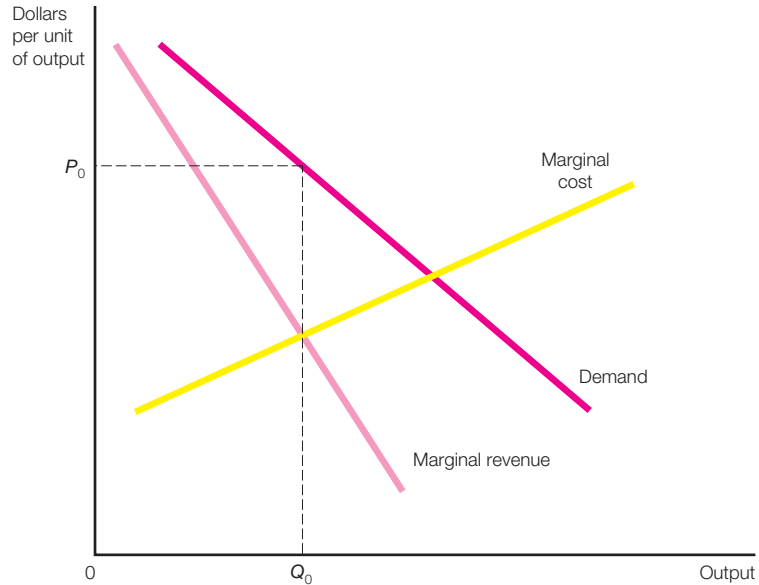
If a cartel is established to set a uniform price for a particular (homogeneous) product, what price will managers charge? To answer this question, managers need to estimate the marginal cost curve for the cartel as a whole. If input prices do not increase as the cartel expands, the marginal cost curve is the horizontal summation of the marginal cost curves of the individual firms. Suppose the resulting marginal cost curve for the cartel is as shown in Figure 11.1. If the demand curve for the industry's product and the relevant marginal revenue curve are as shown there, the output that maximizes the total profit of the cartel members is Q_0 . Therefore, to maximize profit, the cartel will choose a price of P_0 , which is

Cartel When a collusive arrangement is made openly and formally.

FIGURE 11.1

Price and Output Determination by a Cartel

The cartel chooses a price of P_0 and an output of Q_0 .



the monopoly price. Note that this price maximizes the profit earned by the cartel, but it says nothing about how this profit is divided among cartel members.

Cartel managers also determine the distribution of sales across members; this is the process that makes cartels rather unstable. If the purpose of the managers is to maximize the profit to the corporate entity, they should allocate sales to cartel members so that the marginal cost of all members is equal (and, in turn, equal to the cartel's marginal revenue). Otherwise cartel managers can increase corporate profit by reallocating output among members to reduce the cost of producing the cartel's total output. If the marginal cost at firm *A* is higher than that at firm *B*, cartel managers can increase profit by transferring some production from firm *A* to firm *B*.

This allocation of output is unlikely to take place, because allocation decisions are the result of negotiation between members with varying interests and capabilities. This is a political process in which managers have varying amounts of influence. Those with the most influence and the shrewdest negotiators are likely to receive the largest sales quotas, even though this raises the total cartel costs. Also, managers of high-cost firms are likely to receive bigger sales quotas than cost minimization requires because they are usually unwilling to accept the small quotas required by cost minimization. In practice, sales are often distributed in accord

with a member's level of sales in the past or the extent of a member's productive capacity. Also, cartel managers might divide a market geographically, with members being given particular countries or regions.

THE BREAKDOWN OF COLLUSIVE AGREEMENTS

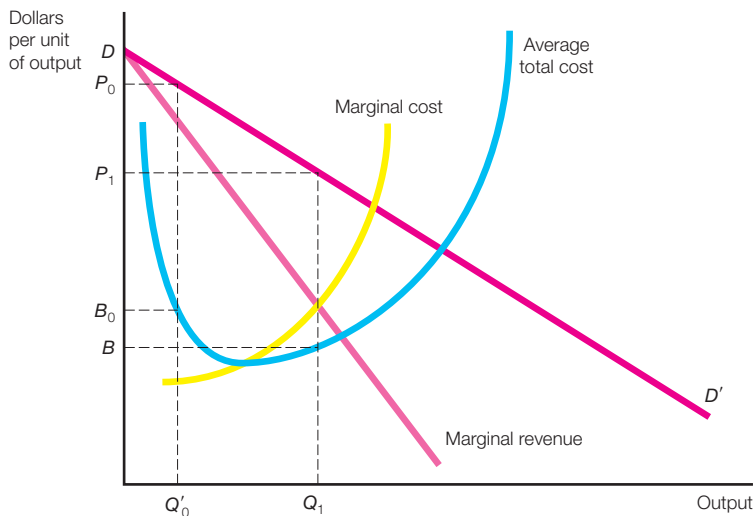
Let's formally see why cartels are not stable so managers understand the weakness of a cartel structure. Consider the firm in Figure 11.2. If firm managers chose to leave the cartel, they would face the demand curve of DD' as long as the other firms in the cartel maintained a price of P_0 . This demand curve is very elastic; managers can significantly expand sales with a small reduction in price. Even if managers were unable to leave the cartel, they would face the same demand curve if they granted secret price concessions.

The maximum profit of managers who either leave the cartel or secretly lower their price is attained when they sell an output of Q_1 at a price of P_1 : This is the output at which marginal cost equals marginal revenue. This price would result in a profit of $Q_1 \times P_1B$, which is generally higher than the profit realized if managers conformed to the price and sales quota dictated by the cartel.² Managers who break away from a cartel—or secretly cheat—increase their profit as long as rival managers do not do the same thing and the cartel does not punish this behavior.

FIGURE 11.2

Instability of Cartels

If the firm leaves the cartel, profit equals $Q_1 \times P_1B$, which is generally higher than it would be if the firm adhered to the price and sales quota established by the cartel.



2. At price P_0 , the firm's profit is $Q'_0 \times P_0B_0$, which is less than $Q_1 \times P_1B$. Because the demand curve is so elastic, total revenues increase significantly as price drops from P_0 to P_1 . While the total costs increase as output increases from Q'_0 to Q_1 , unless they increase very rapidly, the increase in total revenue exceeds the increase in total cost.

But if all managers do this, the cartel disintegrates. Hence there is a constant threat to the existence of a cartel. Its members have an incentive to cheat, and once a few do so, others may follow. Just as is true in relationships, trust is vital to a cartel's existence. As we will see, the incentive to cheat exists even in more informal cooperative endeavors.

PRICE LEADERSHIP

Price leadership In oligopolistic industries, managers at one firm have significant market power and can set their price.

In many oligopolistic industries, managers at one firm have significant market power and can set their price; rivals then follow their lead. This is called the **price leadership** strategy. Examples of industries that have seen the use of this strategy include steel, nonferrous alloys, and agricultural implements. Managers should understand what to consider when setting prices in these markets. We assume the market is composed of a large dominant firm (the price leader) and a number of small firms. Managers at the dominant firm set the price for the market but let the small firms sell all they want at that price. Whatever amount the small firms do *not* supply at that price is provided by the dominant firm.

A new version of price leadership has arisen in the retail sector with the arrival of the “big box” stores. When Wal-Mart or Home Depot come to town, the small retailers, hardware stores, lumber yards, and the like are basically victims of the

STRATEGY SESSION: Cranberries: Where 30% of the Market Are Price Takers

The cranberry, that marvelous red berry that helps prevent bladder infections, has lots of vitamin C, and contains antioxidants, is dominated by a giant growers' cooperative—Ocean Spray.

Ocean Spray is the price setter. When Ocean Spray sets a price (with a 60% share of the U.S. cranberry juice market, 70% share of the fresh cranberry market, 70% share of the cranberry sauce market, and an 80% share of the dried cranberry market), the other (nonmember) producers fall into line. Each year, in late September and early October, Ocean Spray sets a price for sales to supermarkets per case of 24 12-ounce bags. This price is based on anticipated and actual supply and demand conditions in the market.

Given the price that Ocean Spray sets, other producers must decide how much of the product they

wish to harvest for sale, harvest for use in other products (such as juice), or leave in the bogs.

Thus, Ocean Spray is the price leader, and the remaining 30% of the cranberry growers are followers. Whatever price is set by Ocean Spray, the followers take as given and optimize against that price.

Source: Amanda Hesser, “The Case of the Vanishing Berries,” *New York Times*, at www.nytimes.com/2000/11/22/dining/the-case-of-the-vanishing-berries.html; William A. Knudson, “The Economic Impact of Expanded Cranberry Production,” The Strategic Marketing Institute Working Paper 01-1208, Michigan State University, circa 2006–2007 and the current website of the Oppenheimer Group, a distributor of cranberries in the U.S. and Canada, at www.oppyproduce.com/print2.cfm?page=cranberries_var, which states that Ocean Spray has 70% share.

prices charged by those big stores. Small stores may try to differentiate with service and high-end items; but anyone who was selling the items sold by Wal-Mart and Home Depot before their arrival must follow the prices of the big guys. Some do and survive, but newspapers are full of stories of the demise of small businesses that cannot compete with the prices and variety of the big box stores.

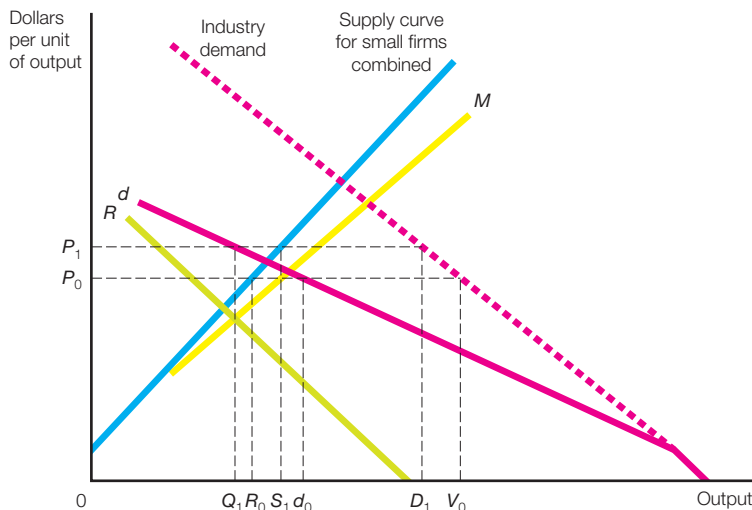
Managers of a dominant firm can readily determine what price to set. Because managers at the small firms are price takers, they act as if they are in a competitive market, taking the price as given. Hence managers at the small firms should choose output where price equals marginal cost. Therefore a supply curve for all the small firms combined is estimated by *horizontally* summing their marginal cost curves. This supply curve is shown in Figure 11.3. The demand curve facing managers at the dominant firm is derived by subtracting the amount supplied by the small firms at each price from the total amount demanded. Thus the demand curve for the output of the dominant firm, d , is determined by finding the *horizontal* difference at each price between the industry demand curve and the supply curve for all small firms combined.

To illustrate how d is derived, suppose managers at the dominant firm set a price of P_0 . The small firms supply R_0 , and the total amount demanded is V_0 . Therefore the output supplied by the dominant firm is $V_0 - R_0$, which is the quantity d_0 on the d curve at price P_0 . In other words, d_0 is set equal to $V_0 - R_0$.

FIGURE 11.3

Price Leadership by a Dominant Firm

Managers at the dominant firm set a price P_1 and supply Q_1 units of the product. The total industry output is D_1 .



PROBLEM SOLVED: Ghoshal, Inc.: A Numerical Example

To illustrate how managers at a dominant firm can determine the price to maximize profit, consider Ghoshal, Inc., the dominant firm in an industry. The demand curve for this industry's product is

$$Q = 100 - 5P$$

where Q is the quantity demanded and P is the price. The supply curve for the small firms in this industry is

$$Q_S = 10 + P$$

where Q_S is the total amount supplied by all these small firms combined. Ghoshal's marginal cost is

$$MC = (8/3)Q_A \quad (11.1)$$

where Q_A is Ghoshal's output.

To derive the demand curve for Ghoshal's output, we subtract Q_S from Q , the result being

$$\begin{aligned} Q_A &= Q - Q_S = [100 - 5P] - [10 + P] = 90 - 6P \\ P &= 15 - (1/6)Q_A \end{aligned} \quad (11.2)$$

Remembering that Ghoshal's total revenue equals PQ_A , Ghoshal's total revenue equals

$$TR = [15 - (1/6)Q_A]Q_A = 15Q_A - (1/6)Q_A^2$$

Therefore, Ghoshal's marginal revenue is

$$\Delta TR_A / \Delta Q_A = MR_A = 15 - (1/3)Q_A \quad (11.3)$$

To maximize profit, managers at Ghoshal need to produce where marginal revenue in equation (11.3) is equal to marginal cost in equation (11.1):

$$MR_A = 15 - (1/3)Q_A = (8/3)Q_A = MC_A$$

So Q_A must equal 5. Consequently, from equation (11.2), it follows that $P = \$14.17$.

To sum up, if managers at Ghoshal want to maximize profit, they should set their price at \$14.17.

QUANT OPTION

Ghoshal's marginal revenue is

$$dTR_A/dQ_A = 15 - (1/3)Q_A$$

The process by which the other points on the d curve are determined is exactly the same; this procedure is repeated at various price levels.

Knowing the demand curve for the output of the dominant firm, d , and the dominant firm's marginal cost curve, M , managers can readily determine the price and output that maximizes their profit. Their marginal revenue curve, R , is derived from the dominant firm's demand curve, d , in the usual way. The optimal output for the dominant firm is the output Q_1 , where its marginal cost equals its marginal revenue. This output is achieved if managers set a price of P_1 . The total industry output is D_1 , and the small firms supply $S_1 (= D_1 - Q_1)$.

POSSIBLE BEHAVIOR IN MARKETS WITH FEW RIVALS

As we will see in the next chapter, managers need to anticipate the behavior of others. As the behavior of rivals changes your actions often change. We now examine models of probable behavior among managers of a small number of firms. For clarity, we will explain behavior using two firms. These markets are known as *duopolies*. As we subsequently show, the results are generalizable to larger oligopolistic markets.

The two firms produce an identical product, and managers make their output decisions simultaneously. When rival managers make decisions without knowing the decisions of others, we say decision making is simultaneous. (We will formally define *simultaneous behavior* and *sequential behavior* in the next chapter.) Managers often make simultaneous decisions. When managers engage in a sealed bid auction they make decisions simultaneously. When firms enter the market at the same time and managers design plant capacity without knowing the plans of others, that too is simultaneous.

Later we consider sequential move strategies, where managers know the decisions of others before making their decisions. Managers who take action before others are called *first movers* or *market leaders*. The market leader is able to accelerate before others for several reasons. First movers often see what others don't because of business acumen, because they invent or patent a product or process, or because entrepreneurs see opportunities others do not. It may also be due to luck: Mobile phone franchises were first granted to companies via a government lottery. Also, landing slots at some airports are also allocated using a lottery.

When Rivals Are Few: Price Competition

One common strategy used by managers is that of price competition. Often price competition results in a downward spiral of price cuts, stopped only (sometimes) by the constraint of marginal cost. The great strategist Sun Tzu referred to behavior such as price wars as that of the scorched earth. Managers should try to avoid this behavior.

Consider two firms, *A* and *B*, producing identical products in a simultaneous-move scenario. Suppose both firms have identical total cost functions of

$$TC_i = 500 + 4q_i + 0.5q_i^2 \quad (11.4)$$

where $i = A, B$ and q_i is the output of firm i .

The market demand as seen by managers of both firms is

$$P = 100 - Q = 100 - q_A - q_B \quad (11.5)$$

where P is the unit price of the product, Q is the quantity demanded at price P , and $Q = q_A + q_B$.

The marginal cost of firm i is

$$MC_i = \Delta TC_i / \Delta q_i = 4 + q_i \quad (11.6)$$

STRATEGY SESSION: Warren Buffett and Duopoly

Why did the “sage” of Omaha make Berkshire Hathaway the largest shareholder in the rating agency Moody’s? According to testimony Buffett gave to the Financial Crisis Inquiry Commission: “The long term value was . . . the duopoly’s . . . incredible pricing power.” The other player in this duopoly is, of course, Standard and Poor’s. If the ratings agencies make mistakes when rating bonds (in particular, giving them higher ratings than they deserve), Buffett’s insurance companies can prosper. For instance, if Buffett’s companies get a higher rating, they can borrow money at a lower interest rate helping them on the cost side and attract customers who are searching for a strong, less risky insurance company on the demand side.

The pricing power works like this: States regulate insurance companies, and states specify which rating agencies can rate the insurance company’s bonds. They choose from a list generated by the Securities and Exchange Commission showing nationally designated rating agencies. Buffett does not have the option of going to multiple agencies for competitive bids but only to those specified by the insurance regulator, and those agencies are usually Moody’s and/or Standard & Poor’s and a few others. But Buffett believes that competition would not produce quality ratings, but rather a race to the bottom. He believes that the agencies would compete not only on price

but also on laxity, that is, come with us and we won’t be tough in giving you a risk rating, but also that a monopoly would have no reason to compete on price or laxity. Thus Buffett likes the duopoly case because of the profits it generates for Moody’s. It turns out that he likes the laxity case too. He states: “What we really hope for is mis-rated securities, because that will give us a chance to make a profit, if we disagree with how the agencies rate them.” Of course, there can be legitimate differences in opinion on ratings, but if the profit comes about because a lax agency rates a skunk as a rose, then questions should be asked. That is what the Commission was doing.

From the text, we know that monopolies don’t have to compete on price—they determine the price, but there is no reason for them to be strict in the quality of their product (because they have no competition). We’ve shown that a duopoly could be very competitive and set prices as if they were in a competitive market. But we’ve also shown that they can behave without colluding and obtain prices and profits close to monopoly results. It is the monopoly end of the spectrum that drove Buffett to Moody’s. The two dominant rating agencies have apparently learned to play our duopoly game well.

Source: John Carney, “Warren Buffett’s Anti-Competitive Profits,” CNBC.com, June 3, 2010, at www.cnbc.com/id/37493375/.

QUANT OPTION

The firm’s marginal cost is

$$dTC_i/dq_i = 4 + q_i$$

If managers at both firms want to compete on price, the competition will drive price down to the level of their marginal cost. Managers should never price below marginal cost because the additional revenue made from making the last sale is exceeded by the additional cost of making that sale.

The first good demanded has a reservation price of \$99—that is, $100 - 1$. This good would cost each producer \$5 (that is, $4 + 1$) to produce. If managers compete on price over this customer, while they could charge as much as \$99, they are willing to accept as little as \$5. In fact, if there were just this one customer, and they competed on price, we expect the price to be \$5. Suppose that managers at firm *A* offer to sell to the customer at \$99 (hoping for a profit of $99 - 5 = 94$). Managers at firm *B* would then offer to sell the product for \$98 (hoping for a profit of $98 - 5 = 93$). Then managers at *A* would counteroffer at \$97. As you can see, in a series of counteroffers, the price will be bid down to \$5. At that price the possibility of a profitable sale by a lower price disappears (any lower price is exceeded by the marginal cost of production).

So we can expect that price equal to marginal cost is the ultimate resolution of this pricing contest. For firm *A*, this means that

$$\begin{aligned} P &= 100 - Q = 100 - q_A - q_B = 4 + q_A = MC_A \\ 2q_A &= 96 - q_B \\ q_A &= 48 - 0.5q_B \end{aligned} \quad (11.7)$$

For firm *B*, it means that

$$\begin{aligned} P &= 100 - Q = 100 - q_A - q_B = 4 + q_B = MC_B \\ 2q_B &= 96 - q_A \\ q_B &= 48 - 0.5q_A \end{aligned} \quad (11.8)$$

We determine the ultimate result of this pricing game by substituting equation (11.8) into equation (11.7) and solving for q_A (or by substituting equation (11.7) into equation (11.8) and solving for q_B)

$$\begin{aligned} q_A &= 48 - 0.5(48 - 0.5q_A) = 24 + 0.25q_A \\ 0.75q_A &= 24 \\ q_A &= 32 \end{aligned}$$

Substituting $q_A = 32$ into equation (11.8) gives $q_B = 48 - 0.5(32) = 32$, which should not be surprising (because each firm has identical cost functions). Because each firm produces 32, the total output is 64 ($= Q$). Substituting $Q = 64$ into the market demand function yields a price of \$36 ($= 100 - 64$). This also equals each firm's marginal cost, \$36 (that is, $4 + 32$). Each firm has a total revenue of \$1,152 (36×32) and total costs of \$1,140 ($= 500 + 4(32) + 0.5(32^2)$), leaving managers at each firm with a profit of \$12.³

3. If each was a constant-cost firm with an identical marginal cost, price competition would take the price down to that marginal cost and profit would be zero (or minus fixed costs if fixed costs exist). For example, if both firms had a marginal cost of \$4, both firms together would produce 96—that is, $4 = 100 - Q$ or $Q = 96$ (with each producing 48). Each firm's total revenue would be \$192 ($= 4 \times 48$), and both firms would have variable cost of \$192 ($= 4 \times 48$). Thus all profit would be lost to competition.

STRATEGY SESSION: Cartels Come in Many Shapes and Sizes

In 1997, every two weeks in Rutherford, New Jersey, 20 shipping line managers met and discussed what they would charge to move cargo across the North Atlantic. Although this meeting could have been composed of the managers of a single ocean shipping line discussing pricing strategy, it was not. Rather, they were executives of 20 different companies. Exempt from the U.S. antitrust laws, they collusively set rates on tens of billions of dollars of cargo. This practice has since been declared illegal, but it demonstrates that, if allowed, many firms would choose to collude rather than to compete.

Their monopoly power was limited because the cartel was unable to control the shipping capacity of its members, and some ocean carriers were not members of the cartel. Nevertheless, it was estimated that the cartel was able to raise rates 18 to 19% above competitive rates. Ah, the power of monopoly.

Of course, the epitome of controlling supply and hence creating monopoly power is the Organization of Petroleum Exporting Countries [OPEC] and its power in the crude oil market. Despite the fact that the 12 nations (Algeria, Angola, Ecuador, Iran, Iraq, Kuwait, Libya, Nigeria, Qatar, Saudi Arabia, United Arab

Emirates, and Venezuela) that make up OPEC supply only approximately 40% of the world's oil production, OPEC dictates the world's price. Other oil producing nations have chosen not to join and/or their antitrust laws prohibit firms from their countries from joining. This remaining 60% of the market is content with OPEC marking prices up considerably over marginal costs (because it yields these countries or firms in these countries handsome profits).

The monopoly power is so substantial because of the strong demand for the product, the relative inelasticity of demand for the product, and the fact that the low-cost producers are members of the cartel. Although the cartel is not perfect (that would be where all producers are members and all their capacity is controlled), the nonmembers are high cost producers (which makes it difficult for them to challenge the cartel—since the cartel could take down prices to drive the high cost producers out of business).

In 2011, there was substantial unrest in the Middle East. While this unrest has substantially impacted Libyan oil production (virtually zero production in summer 2011), as of this writing there has been no impact on production in Algeria (where protests against the

4. It is the horizontal summation because we want to measure the additional cost of producing an additional unit of output in the cartel. To produce that additional unit in the cheapest possible way, the cartel would always want to have identical marginal costs for each producer. If the marginal costs were not the same, the cartel could lower its total cost of production by shifting production from the high-marginal-cost firm to the low-marginal-cost firm until their marginal costs were equalized. See the appendix to Chapter 8 and the earlier discussion in this chapter.

When Rivals Are Few: Collusion

What if our two managers both realized the dangers of a price war and instead chose to act cooperatively? How cooperative could they get? Let's take it to the limit and assume they can legally form a cartel. Under these circumstances, the market demand curve is the cartel's demand curve, and the cartel's marginal cost curve is the *horizontal* summation of each firm's marginal cost curve.⁴ Rewriting each firm's marginal cost as

$$\begin{aligned}q_A &= -4 + MC_A \\ q_B &= -4 + MC_B\end{aligned}$$

and adding q_A and q_B (that is, summing the marginal costs horizontally—adding up the quantities produced at any given marginal cost), we get

government fizzled with announced plans for democratic reforms), Kuwait (where the emir gave people free food and a \$4,000 cash grant), Qatar (where things quieted down after calls for protest against the emir), Saudi Arabia (where protests quickly subsided), and the United Arab Emirates (where 160 intellectuals petitioned for comprehensive reform). Of course, given the ongoing activities in Iraq, production there is subject to uncertainty. Most of the remaining countries, such as Nigeria, are not noted for their stability.

But the crisis in the Middle East (with the exception of Libya)—in countries like Egypt, Tunisia, Syria, and Yemen—basically involves countries that are not members of OPEC and are not significant oil producers.

Others have tried to emulate the behavior of OPEC. Brazil and Colombia (both of which have substantial oil production but are not members of OPEC) once looked at their neighbor Venezuela with envy. Venezuela was able to receive a very high price (because of OPEC) relative to the cost of extraction for a black liquid that was turned into energy. Brazil and Colombia noted that they had a product that could be turned into another black liquid that produced energy too—not oil, but coffee. So they formed a coffee cartel.

Ever heard of it? Not likely. It failed miserably. They withheld coffee from the market to raise price (just like cartel theory dictates), but when the price rose, consumers switched from coffee to tea and caffeinated colas. Though managers may successfully control supply in a market, they may not be able to exercise their monopoly pricing power if close substitutes exist. Close substitutes dictate that demand is relatively elastic, that is, consumers are price sensitive.

Cartels exist not only for goods, but also for services. In Germany, a wage-setting cartel is constitutionally sanctioned between unions and corporations. The cartel keeps wages high and labor strife low, and ends up costing Germans jobs. This is exactly what we expect a monopoly to do—raise prices (wages) and restrict output (jobs).

Sources: "As U.S. Trade Grows, Shipping Cartels Get a Bit More Scrutiny," *Wall Street Journal*, October 7, 1997; and "German Wage Pact Ends Up Costing Jobs," *Wall Street Journal*, February 19, 1997.

$$Q = q_A + q_B = -4 + MC_A - 4 + MC_B = -8 + 2MC$$

Rearranging yields

$$\begin{aligned} 2MC &= 8 + Q \\ MC &= 4 + 0.5Q \end{aligned}$$

which is the cartel's marginal cost.

The cartel behaves as a monopolist (see Chapter 8 and earlier in this chapter) and sets its marginal revenue equal to its marginal cost. The cartel's total revenue is

$$TR = PQ = (100 - Q)Q = 100Q - Q^2$$

The cartel's marginal revenue is therefore

$$MR = \Delta TR / \Delta Q = 100 - 2Q$$

STRATEGY SESSION: How Seagoing Chemical Haulers May Have Tried to Share the Market

In 2003 Stolt-Nielsen SA, two of its executives, and two subsidiary companies were accused of conspiring to allocate customers, fix prices, and rig bids on contracts in the parcel tanker business. It was alleged that the participants would allocate business on a route-by-route basis, and that Stolt-Nielsen had calculated a table that showed how much better off their revenues would be if they cooperated rather than competed.^a This accusation occurred after Stolt-Nielsen had gone to the U.S. Department of Justice (DOJ) on January 8, 2003, and confessed to the antitrust violation to which it had been a party. Soon thereafter, the Justice Department entered into a leniency agreement (the amnesty agreement outlined below) with Stolt-Nielsen.

However, on April 8, 2003, the Justice Department told Stolt that it had learned that Stolt-Nielsen had not desisted from illegal activity, and so it arrested the manager of Stolt-Nielsen's tanker trading division. In February 2004, Stolt and the manager brought a civil action suit against the Department of

Justice seeking an end to its prosecution. On January 14, 2005, the district court in the Eastern District of Pennsylvania enjoined the Justice Department from terminating Stolt-Nielsen's immunity. The Third Circuit Court of Appeals, however, upheld the ruling in favor of the DOJ. But there's more. The appeal court said that if Stolt-Nielsen asserted the leniency agreement as a defense *after indictment*, then the reviewing court had to consider the agreement anew, determine the date Stolt-Nielsen discovered the conspiracy, consider Stolt-Nielsen's subsequent actions, and decide whether Stolt took prompt and effective action to end their role in the illegal activity. Stolt-Nielsen then asserted the leniency agreement as a defense. After viewing Stolt-Nielsen's actions regarding the above, the Third Circuit ruled that Stolt-Nielsen complied with what the appeal court wanted and dismissed the indictments. The Third Circuit's message was clear: "a deal is a deal" and the Justice Department cannot renege unless it has extraordinary reasons for doing so.^b

QUANT OPTION

The cartel's marginal revenue is

$$dTR/dQ = 100 - 2Q$$

Setting the cartel's marginal revenue equal to the cartel's marginal cost yields

$$\begin{aligned} MR &= 100 - 2Q = 4 + 0.5Q = MC \\ 2.5Q &= 96 \\ Q &= 38.4 \end{aligned}$$

Substituting $Q = 38.4$ into the cartel's demand curve gives a price of \$61.6 (that is, $100 - 38.4$). The cartel's total revenue is \$2,365.44 (that is, $\$61.6 \times 38.4$). Since

Here's a little background. The initial indictment was a surprise because Stolt-Nielsen had entered into an amnesty agreement with the Justice Department's Antitrust Division. Prior to the indictment, Stolt-Nielsen claimed it had ceased its anticompetitive actions and had provided information helpful in the convictions of its co-conspirators (which entailed over \$62 million in fines for Odfjell Seachem and Jo Tankers, and jail time for some of their executives). The surprise came because the Justice Department revoked the amnesty agreement. The amnesty program was 13 years old in 2006, and Stolt-Nielsen would be the first revocation. The agreement is designed to encourage whistleblowers to come forward and provide supporting information. This is a first-mover situation: The first firm that comes forth gets the amnesty if it terminates the criminal activity and provides information for the prosecution.

As a result of this case, the Justice Department issued a new leniency letter for firms seeking amnesty for antitrust violations. The Justice Department

has pledged to increase the transparency and predictability of its amnesty program. These changes entail a burden of proof on the amnesty applicant to prove that it has stopped all illegal activity. In addition, the leniency applicant must waive judicial review or pre-indictment leniency revocations.^c

The amnesty/leniency agreement is an important weapon in antitrust law. It creates an incentive for a firm or an individual to come forward and "spill the beans." Only one firm/individual gets a bite at the apple, so if something is amiss, there is an incentive to move fast. Moving second means prosecution. There is only one winner in this race.

^a See James Bandle, "How Seagoing Chemical Carriers May Have Tried to Divide Market," *Wall Street Journal*, February 20, 2003.

^b See the website of the National Association of Criminal Defense Lawyers, at www.nacdl.org/public.nsf/PrinterFriendly/A0807p26?

^c See "Department of Justice Issues New Antitrust Leniency Letter," December 11, 2008, at www.pepperlaw.com/publications_update.aspx?ArticleKey=1319.

each firm has the same marginal cost equation, each should produce the same amount, 19.2, so that both have a marginal cost of \$23.2 (that is, $4 + 19.2$), which, of course, equals the cartel's marginal revenue of 23.2 (that is, $100 - 2(38.4)$). The two firms split the total revenue so that each receives \$1,182.72. Each firm has a total cost of \$761.12 ($= 500 + 4(19.2) + 0.5(19.2^2)$); hence, each firm makes a profit of \$421.6, a considerable improvement over the \$12 made when the firms competed on price. Note the cooperative behavior significantly restricts output (from 64 to 38.4) and significantly increases price (from \$36 to \$61.6); but as observed in Chapter 8, that is what monopolists do.

When Rivals Are Few: Quantity (Capacity) Competition

Unfortunately, forming a cartel is often illegal. But, strictly competing on price is a lose-lose strategy. So, what are managers to do? Well, they could compete on something other than price. Managers should try to compete on any metric that affects profit and gives them a higher profit relative to competing on price. One metric

that jumps to mind is quantity (or production capacity). This type of competition is sometimes called *Cournot*, named after the French economist who initially derived its properties.

Cournot analysis makes the following assumptions: The rival managers move simultaneously, have the same view of the market demand, estimate each other's cost functions, and choose their profit-maximizing output conditional on their rival choosing the same.

It is thought as problematic that rival managers hold the same beliefs regarding demand. Yet, there is evidence that it is more likely to occur than students think. In many situations, government- or trade-association-generated data (macroeconomic and industry-specific variables) are used by all analysts and the corporate intelligence of each firm observes the same economic landscape. Often, managers compete against one another for long periods of time. They often get to know each other through various professional associations. Nevertheless, it is possible for two (or more) managers viewing the same economic data to come up with different conclusions or assessments. With respect to knowing each other's cost functions, in some cases, good approximations of adversaries' costs are likely. For instance, in the airline industry, there are only two manufacturers of large aircraft remaining (Boeing and Airbus) and only several producers of small commercial aircraft (for example, Bombardier and Embraer). The carriers are either flying the same aircraft as their adversaries or have "spec'd" the aircraft (that is, received all the operating characteristics of that aircraft from the manufacturer when the carrier considered purchasing new aircraft). Personnel are unionized (for the most part) and wage rates are well known. All carriers buy fuel, food, and other items from a limited set of suppliers. Therefore, carrier *A* has a fairly decent estimate of what it costs carrier *B* to operate its fleet (and vice versa). In addition, executive talent is mobile within the industry, and when executives leave carrier *A*, they carry knowledge of the company in their heads that can be useful to company *B*.⁵

The last assumption is that firm *A* optimizes their quantity (capacity) given that firm *B*'s quantity (capacity) is fixed. This is not as restrictive as it sounds. We first formulate this situation as a series of "what-if" questions. *If* my adversary actually is going to produce quantity *X*, *what* quantity would I produce to maximize my profit? Which output you actually choose to produce of all the "what-if" possibilities depends on what you think your adversary will *actually* do (and your adversary is going through the same "what-if" process). By a process of deduction, managers can estimate the most logical output for rivals given profit-maximizing behavior. This is the Cournot solution, and we see that it yields a solution identical to that of game theory (see Chapter 12).

We now view the Cournot solution to the preceding case in two different ways. The first is by following a series of "what-if" scenarios. We deal with the decisions of managers at firm *A* and treat this as our firm. Clearly, if firm *A* managers think that managers at firm *B* will abdicate the market to them, they should

5. We will leave it to your ethics course to debate about what information can be revealed. To see how information is obtained about competitors, see "They Snoop to Conquer," *BusinessWeek*, October 28, 1996. In a classic case of crossing the line, Jose Ignacio Lopez left General Motors in 1992 allegedly carrying a briefcase of GM blueprints to his new job at Volkswagen. You need no ethics course to know that this is wrong. GM sued, and in 1996 the case was settled. Lopez was forced to resign from VW, and VW was required to pay GM \$100 million and purchase \$1 billion worth of GM auto parts.

behave as a monopolist. Since the monopolist's marginal revenue is the same as the cartel's in the preceding situation (and now firm A 's marginal revenue, since it is the only producer in the market), and firm A 's marginal cost is $MC_A = 4 + q_A$, managers maximize profit by setting $MR = MC$ or

$$\begin{aligned}MR_A &= 100 - 2q_A = 4 + q_A = MC_A \\3q_A &= 96 \\q_A &= 32\end{aligned}$$

So, if $q_B = 0$, the profit-maximizing, optimal-quantity response of firm- A managers is $q_A = 32$.

On the other hand, if managers think their rival will produce 96 units, they will only produce 4 (at most). Rewrite the market demand curve as $Q = 100 - P$. Firm A 's *residual* demand curve (i.e., the market demand curve less what managers assume firm B produces), what is left for firm A after firm B managers make their production decision, is

$$\begin{aligned}q_A &= 100 - P - 96 = 4 - P \\P &= 4 - q_A\end{aligned}$$

Firm- A managers will produce nothing under these circumstances, since the *highest* the price could be is \$4 (when q_A is zero) and the lowest their marginal cost could be is \$4 (when q_A is zero). Therefore, the price could never be equal to or exceed firm A 's marginal cost at a positive level of output. Therefore, if $q_B = 96$, the profit-maximizing, optimal-quantity response of firm- A managers is $q_A = 0$.

Suppose firm- A managers think their rival will produce 50 units; that is, $q_B = 50$? Under these circumstances, the residual demand is

$$\begin{aligned}q_A &= 100 - P - 50 = 50 - P \\P &= 50 - q_A\end{aligned}$$

Firm A 's total revenue is $Pq_A = (50 - q_A)q_A = 50q_A - q_A^2$. Hence, marginal revenue is

$$MR_A = \Delta TR_A / \Delta q_A = 50 - 2q_A$$

QUANT OPTION

The firm's marginal revenue is

$$dTR/dq_A = 50 - 2q_A$$

TABLE 11.1

Profit-Maximizing Output Responses of Managers of Firm A Given Their Assumptions about Firm B Output

If Firm B Produces	Then Firm A Produces
0	32
50	15.33
96	0

To maximize profit, set $MR_A = MC_A$, or

$$MR_A = 50 - 2q_A = 4 + q_A = MC_A$$

$$3q_A = 46$$

$$q_A = 15.33$$

Therefore, if $q_B = 50$, the profit-maximizing, optimal-quantity response of firm-A managers is $q_A = 15.33$.

So, we know the optimal profit-maximizing responses for the “what-if” scenarios we investigated (see Table 11.1).

By doing more “what-if” situations, we can complete Table 11.1 for all possible firm B outputs between 0 and 96. But, we can get the equivalent of a full table directly with the analysis that follows.

Firm A maximizes profit when its total revenue (Pq_A) exceeds its total cost ($500 + 4q_A + 0.5q_A^2$) by the maximal amount. Total revenue is

$$TR = (100 - Q)q_A = (100 - q_A - q_B)q_A = 100q_A - q_A^2 - q_Aq_B$$

Marginal revenue is

$$MR_A = \Delta TR_A / \Delta q_A = 100 - 2q_A - q_B \quad (11.9)$$

QUANT OPTION

The firm's marginal revenue is

$$\partial TR_A / \partial q_A = 100 - 2q_A - q_B$$

To maximize profit, set $MR_A = MC_A$, or

$$MR_A = 100 - 2q_A - q_B = 4 + q_A = MC_A$$

$$3q_A = 96 - q_B$$

$$q_A = 32 - (1/3)q_B \quad (11.10)$$

Equation (11.10) is called firm A's **reaction function**; that is, it identifies for managers the profit-maximizing output to produce *given* the output of rivals at firm B. Every number in the right-hand column of Table 11.1 occurs when the corresponding number on the left-hand side of Table 11.1 is substituted into the reaction function (11.10).

Because firm B has the same cost function as firm A and both face the same market demand curve, firm B's reaction function is

$$q_B = 32 - (1/3)q_A \quad (11.11)$$

We identify how to anticipate the profit maximizing output of a rival by substituting firm A's reaction function (11.10) into firm B's reaction function (11.11) and solving for q_A . Alternatively, we could substitute firm B's reaction function (11.11) into firm A's reaction function (11.10) and solve for q_A . The latter yields

$$\begin{aligned} q_A &= 32 - (1/3)[32 - (1/3)q_A] \\ q_A &= (96/3) - (32/3) + (1/9)q_A \\ (8/9)q_A &= (64/3) \\ q_A &= 24 \end{aligned}$$

Substituting $q_A = 24$ into firm B's reaction function (11.11) gives

$$q_B = 32 - (1/3)24 = 24$$

Therefore, $Q = q_A + q_B = 48$ and substituting $Q = 48$ into the market demand function gives a price of \$52; that is, $100 - 48$.

Think of it like this: The only way that managers at both firms can maximize profit is if they stay on their reaction functions. Again, this function identifies for managers the profit-maximizing output conditional on the output of their rival. The only way this is possible is if we find the point of intersection between the two functions. This then represents the only quantities where what one manager wants to do is mutually consistent with what the other manager wants to do. This is known as a Nash equilibrium after Nobel laureate John Nash—subject of the book and the movie *A Beautiful Mind*.

Under the Cournot scenario, managers at each firm produce 24 units and the market price is \$52. Each firm's total revenue is \$1,248 (i.e., 52×24), and each firm's total cost is \$884 (that is, $500 + 4(24) + 0.5(24^2)$), so managers at each firm earn a profit of \$364. While this is less than each share of the monopoly (cartel) profit of \$421.6, it is considerably better than what they will earn if they compete on price (\$12). The significance of this is the powerful impact of adding just one more equal producer to a monopoly market. In this case, the price falls by 15.6% (from \$61.6 to \$52) and output increases by 25% (from 38.4 to 48). Hence, adding a competitor can have a significant impact on mitigating the power of a monopolist. Looked at from the other direction, the two firms acting

Reaction function A function that identifies for managers the profit-maximizing output to produce given the output of their rivals.

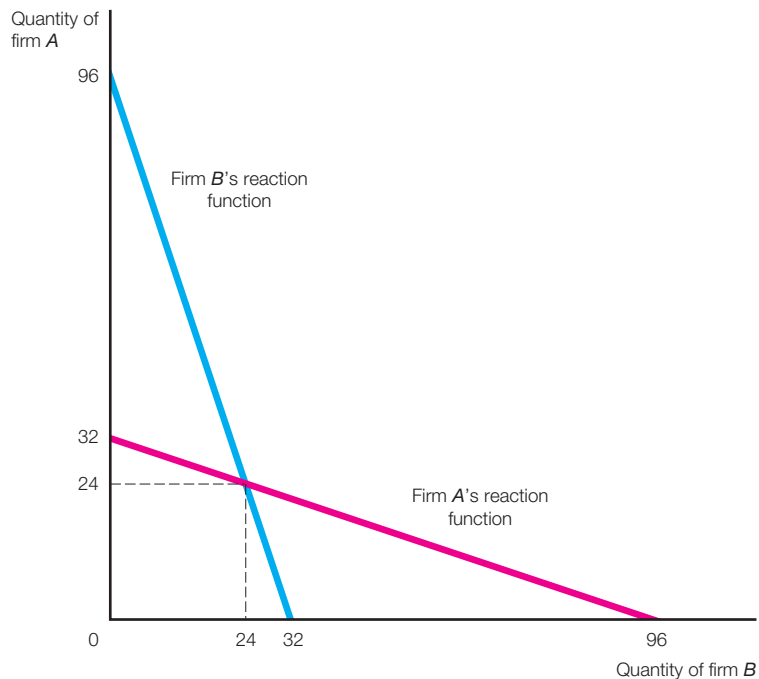
as Cournot quantity competitors can retain 86.3% of the monopoly (cartel) profit and enhance the price-competitive profit slightly over 30-fold.

How can managers get themselves into a Cournot scenario and avoid a price-competitive scenario? Some managers cannot seem to avoid the latter; for example, the airlines constantly seem to engage in price wars, much to the detriment of their profit. On the other hand, managers who learn not to “rock the boat” or to “kill the goose that lays the golden egg” learn to compete on quantity (capacity) and not price. Examples include General Electric and Westinghouse in the marketplace for steam turbine engines and Rockwell International and others in the market for water meters.⁶ As can be seen from the preceding, the stakes are high, so this quantity (capacity) competition is a strategy worth learning for managers who have only a few market rivals. Figure 11.4 is a picture of the situation just discussed.

FIGURE 11.4

Cournot Reaction Functions for Firms A and B

A Cournot equilibrium occurs where the two firms' reaction functions intersect. This is the only output combination where both firms' expectations of what the other firm will produce are consistent with their own expectations of their own optimal output. In this case, both firms produce 24 units.



6. See Nancy Taubenslag, “Rockwell International,” Harvard Business School Case, 9383-019, July 1983.

Such behavior tends to work well when large capital investments must be made. Once the capital investment is made, the quantity to be produced is pretty much determined. Airlines do not practice Cournot competition well because it is easy to change capacity by leasing aircraft (e.g., all of Continental Airlines planes are leased). Oil refining tends to work well because it is difficult to build new refineries (i.e., it's expensive, creates environmental concerns, and no one wants one located near them—the last refinery built in the United States was Marathon-Ashland's Garyville, Louisiana, facility completed in 1976).

The Cournot Scenario with More than Two Firms

Consider a market demand curve $P = a - bQ$ with n identical firms, i.e., $Q_i = Q/n$ where a and b are parameters of the demand function and Q_i is the output of the i th firm. The marginal cost of each firm is $MC_i = c + eQ_i$, where c and e are parameters of the marginal cost function. The market demand curve can be rewritten as

$$P = a - bQ_1 - bQ_2 - \dots - bQ_i - \dots - bQ_n$$

The total revenue for firm i is

$$\begin{aligned} TR_i &= P \times Q_i = (a - bQ_1 - bQ_2 - \dots - bQ_i - \dots - bQ_n)Q_i \\ &= aQ_i - bQ_iQ_1 - bQ_iQ_2 - \dots - bQ_i^2 - \dots - bQ_iQ_n \end{aligned}$$

The marginal revenue of firm i is

$$\begin{aligned} MR_i &= \Delta TR_i / \Delta Q_i = a - bQ_1 - bQ_2 - \dots - 2bQ_i - \dots - bQ_n \\ &= a - 2bQ_i - \sum_{k \neq i}^n bQ_k = a - 2bQ_i - (n-1)bQ_i \end{aligned}$$

since all firms are the same, $Q_i = Q_k$ for $k \neq i$.

QUANT OPTION

So that we do practice what we preach

$$\begin{aligned} MR_i &= \partial TR_i / \partial Q_i = a - bQ_1 - bQ_2 - \dots - 2bQ_i - \dots - bQ_n \\ &= a - 2bQ_i - \sum_{k \neq i}^n bQ_k = a - 2bQ_i - (n-1)bQ_i \end{aligned}$$

Further simplifying the marginal revenue gives $MR_i = a - (n+1)bQ_i$. To maximize profit, managers will set $MR_i = MC_i$ or

$$MR_i = a - (n+1)bQ_i = c + eQ_i = MC_i$$

STRATEGY SESSION: The End of a Duopoly?

At the 2011 Paris Air show, Jim Albaugh, the head of Boeing's civil jet division, stated that Embraer (of Brazil), Bombardier (of Canada), Comac (of China), and Irkut (of Russia) are likely to challenge Boeing and Airbus in the duopoly that the two carriers have had in the 100–200 seat aircraft market for the last 15 years. The latter three companies already have aircraft in the development stages and have secured orders. In both the Chinese and the Russian cases, the orders are from domestic carriers in their countries. Albaugh said, "The days of the duopoly with Airbus are over." This was seconded by Tom Enders, Airbus's chief executive. However, Enders added, "The duopoly is over in the 100 to 150 (seat) aircraft segment because this is where the new entrants . . . want to be . . . so that doesn't mean the duopoly is over in the entire range of products."

The significance of the entry into the 100–200 seat market (even at the lower end) is that such sizes dominate the commercial jet market. In the next 20 years, 70% of the market of approximately 25,000 new jets will be in that size range. Enders doubts

whether the market can sustain six firms and predicts likely consolidation. And let us not forget that India will have the capability of entering the market in the future. It's also difficult to think that the Chinese won't go after the larger jet market after they cut their teeth on the smaller jets.

But what about the market for jets with fewer than 100 seats (where Boeing and Airbus don't play)? This market includes regional jets and turbo props; it's a triopoly in which Bombardier holds 38% of the market, Embraer holds 36%, and ATR holds 26%. Without turboprops, it's a duopoly of Bombardier and Embraer. The market for engines to power these aircraft is a duopoly between General Electric and Pratt and Whitney, with both at about half the market.

Source: Mark Odell, "Boeing and Airbus Call Time on Duopoly," *Financial Times*, June 21, 2011 at www.ft.com/intl/cms/s/55a1fcf0-9b39-11e0-a254-00144feabdc0,dwp_uuid=890484; Max Kingsley-Jones, "Numbers Game: We crunch the data and analyze how the battle between ATR, Bombardier, and Embraer played out in the space below the mainline jet sector during 2010," *Flight International*, March 29, 2011.

Solving for Q_i gives

$$Q_i = (a - c) / [(n + 1)b + e]$$

Table 11.2 shows the situation when multiple firms (each of which is identical to the two Cournot firms) compete on quantity. Using the formula for Q_i above with $a = 100$, $b = 1$, $c = 4$, and $e = 1$ yields $Q_i = 96 / (n + 2)$. Adding a third equal (in cost) competitor to our duopoly drops the price 31.17% from the cartel monopoly price and increases output by 50%. Having eight equal-in-cost competitors in this Cournot situation drops the price by 62.34% relative to the cartel monopoly price and increases output by 100%. Profit becomes negative after three Cournot competitors because of the high level of fixed costs (\$500). If the fixed cost were lower (say \$50), all situations depicted (except for high values of n) would entail positive profit.

TABLE 11.2

Price, Output, and Profits with Multiple Cournot Competitors

Number of Competitors	Price	Percentage Decrease	Quantity/ Firm	Profit/ Firm	Total Quantity	Percentage Increase
Cartel	61.6		19.2	421.6	32	
2	52	15.58	24	364	48	25
3	42.4	31.17	19.2	52.96	57.6	50
4	36	41.56	16	-116	64	66.67
5	31.43	48.98	13.71	-217.88	68.57	78.57
6	28	54.55	12	-284	72	87.50
7	25.33	58.87	10.67	-329.33	74.67	94.44
8	23.2	62.34	9.6	-361.76	76.8	100
9	21.45	65.17	8.73	-385.75	78.55	104.55
10	20	67.53	8	-404	80	108.33
n	$\frac{4n + 200}{n + 2}$		$\frac{96}{n + 2}$	$\frac{(11,824 - 2,000n - 500n^2)}{(n + 2)^2}$	$\frac{96n}{n + 2}$	
∞	4	93.51	0	-500	96	150

This analysis shows just how the addition of a few entrants in a Cournot situation can bring significant price competition into the market. Even if managers of these entrants have higher costs, they still erode the market power of incumbents and generate significant downward pressure on price.

When Managers Move First: Stackelberg Behavior

Now consider a situation in which managers at one firm are able to implement actions prior to those of rival managers. For example, managers at firm *A* choose and credibly commit to a capacity decision; managers at firm *B* know the decision when they choose their own capacity. One behavior is described by Heinrich Freiherr von Stackelberg, and is named after him. How should managers at firm *B* react to the capacity decision of *A* managers? Well, if they want to maximize profit, they have to follow their reaction function. This function represents how *B* managers should act to maximize profit, given the decision of *A* managers. Managers cannot possibly maximize profit if they are operating off their reaction function. So, managers at firm *A* can anticipate the capacity choice of *B* managers.

Therefore, the demand curve firm *A* faces reads (after substituting firm *B*'s reaction function, equation (11.11), for q_B in the market demand curve, since we can anticipate the decision of *B* managers)

$$P = 100 - q_A - q_B = 100 - q_A - [32 - (1/3)q_A] = 68 - (2/3)q_A$$

STRATEGY SESSION: Competition among the Few

Below are three related stories about markets in which there is limited competition. Taken together they show that while more competitive markets benefit the consumer, a greater market concentration also carries benefits. The first two stories are about small competitors trying to enter the market and increase competition and their entry strategies when big firms dominate the market.^a In the third story, competition is being eliminated by a merger that will create a firm with a market share approaching 70% in the Philippine long-distance market.

The world's largest accountancy firms (Price-waterhouseCoopers, Ernst & Young, Deloitte, and KMPG) account for about 90% of the world's large company audit market. There are also smaller rivals clamoring for a chance to increase their participation. They raise issues that a big failure can take out one of the big firms (think Enron taking out Arthur Andersen). Managers at these smaller firms recommend several strategies to make the market more competitive. These include joint auditing (which the big players say will raise the cost of an audit by 20%)

and a mandatory rotating of audits. The pressure is on. The European Union issued a report titled *Audit Policy: Lessons from the Crisis*, and the small firms are hopeful that recommendations (scheduled for November 2011) will open up the market for them.

The second case of increasing competition in markets was discussed in the Strategy Session box on page 418 where we referenced, Warren Buffett's investment in Moody's and his statement that it was a duopoly. While Moody's and Standard & Poor's each have about 40% of the market, a third player, Fitch's, has about 20%. The article refers to the industry as an "oligopoly." Another player, Jules Kroll, wants to enter the market. He previously ran a firm, sold it for \$1.9 billion, and started another doing the same thing as the first—sleuthing for one firm on another firm. He feels that he can enter the market because the three major ratings agencies lost their credibility by not performing due diligence on mortgage-backed securities and other structured investments prior to the financial crisis of 2007. He pledges to "look under the covers" when he generates ratings, but he

Firm A's total revenue is $Pq_A = [68 - (2/3)q_A]q_A = 68q_A - (2/3)q_A^2$. The firm's marginal revenue is

$$MR_A = \Delta TR_A / \Delta q_A = 68 - (4/3)q_A$$

QUANT OPTION

Firm A's marginal revenue is

$$dTR_A/dq_A = 68 - (4/3)q_A$$

accepts the historical model in which the issuers of the securities pay the rating agencies. This creates a moral hazard of the firm pandering to the issuer to get the business of rating the issuer's securities. Kroll's target is a 10% market share in five years. The big guys aren't particularly worried, pointing out Kroll's lack of geographical coverage and industry expertise relative to their positions.

Our third case involves the Philippine Long Distance Telephone Company which has acquired Digitel Telecommunications Philippines, resulting in what is being called a duopoly in the wireless segment. So says Justino B. Calaycay, Jr., an analyst at Accord Capital Equities Corporation. The merger will result in the firm controlling 60–70% of the wireless business. Globe Telecom is the other major player. Cid L. Terosa, senior economist at the University of Asia and the Pacific, said the merger had some potentially good points and some bad points. The good points related to economies of scale. The bad points are potential collusive behavior. Calaycay noted that if the two duopolists price competed for consumers'

"peso-votes," it would be good for the consumer. However, if one player got too big, "it might eventually result in a virtual monopoly." Astro C. del Castillo, managing director of the brokerage firm First Grade Holdings, has stated that "Duopoly is an advantage in the Philippines, since it will instill healthy competition between the two dominant firms, as they will compete to provide the most economical and beneficial telecom service for the consumers."

^a 90% in the case of the top four accountancy firms, and almost 100% in the case of the top three security rating firms—note that usually an 80% share for four firms is the benchmark for very high concentration.

Sources: "Four for All, All for Four," February 25, 2011; *Financial Mail* (South Africa); Janet Morrissey, "A Corporate Sleuth's Eye Turns to Financial Ratings; Experienced Investigator Sees Chance to Compete with the Likes of Moody's," *The International Herald Tribune*, February 28, 2011; J. D. T. Chua, "Popular Economics: When Only Two Firms Control the Market," *Business World* (Philippines), April 26, 2011.

Managers at firm A set marginal revenue equal to marginal cost to maximize profit

$$\begin{aligned}MR_A &= 68 - (4/3)q_A = 4 + q_A = MC_A \\(7/3)q_A &= 64 \\q_A &= 27.43\end{aligned}$$

Substituting $q_A = 27.43$ into firm B's reaction function yields

$$q_B = 32 - (1/3)27.43 = 22.86$$

Therefore, $Q = q_A + q_B = 50.29$ and substituting $Q = 50.29$ into the market demand curve gives a price of \$49.71 ($100 - 50.29$). Firm A's total revenue is \$1,363.59 ($\49.71×27.43), and firm A's total cost is \$985.88 [$500 + 4(27.43) + 0.5(27.43^2)$]; therefore, managers earn a profit of \$377.71 (which is \$13.71 better than the simultaneous decision of Cournot). As a general managerial

rule, if you have the market strength so market rivals cede you the power to move first, use it. Firm B 's total revenue is \$1,136.33 ($\49.71×22.86), and its total cost is \$852.65 (that is, $500 + 4(22.86) + 0.5(22.86^2)$); therefore, the profit is \$283.67 (which is \$80.33 worse than under Cournot, so, managers at B do pay a penalty for moving second).

The profit situation is exactly reversed if managers at firm B moved first. In this case, where the firms have the same costs, it is worth the same amount for each firm to go first, \$94.04 (i.e., the gain from going first plus the loss if the firm moves second). In situations where the firms have different cost functions, the low-cost firm has a greater advantage than the high-cost firm in all the pricing schemes discussed here, including the first-mover situation. Managers at a low-cost firm have the most to gain by moving first. They can even afford to "purchase" the first-mover advantage, outbid the high-cost firm for the patent on the product, or build a bigger plant than the high-cost firm to preempt its output decision.

Let us see how profit changes when managers face different cost functions. The two firms still face the demand curve $P = 100 - q_A - q_B$. But, now, managers at firm A face a cost function $TC_A = 500 + 4q_A + 0.5q_A^2$, while managers at firm B face one of $TC_B = 500 + 10q_B + 0.5q_B^2$ (i.e., firm B has higher costs than firm A). Firm A 's reaction function is $q_A = 32 - (1/3)q_B$, and firm B 's reaction function is $q_B = 30 - (1/3)q_A$. If you solve for the Stackelberg solution with managers at firm A choosing first, $P = \$51.143$, $q_A = 28.286$, $q_B = 20.571$, $\pi_A = \$433.43$, and $\pi_B = \$134.78$. If you solve for a Stackelberg solution with firm B going first, $P = \$51.43$, $q_A = 23.714$, $q_B = 24.857$, $\pi_A = \$343.55$, and $\pi_B = \$220.86$.

Now we can illustrate how lower costs leverage the advantage of moving first. If managers at firm A move first, they earn $\pi_A = \$433.43$, whereas if they wait for firm B managers to move first, they earn $\pi_A = \$343.55$. Therefore, managers at firm A gain $\$433.43 - \$343.55 = \$89.88$ by moving first. If managers at firm B move first, they get $\pi_B = \$220.86$, whereas if they go second, they earn $\pi_B = \$134.78$. Therefore, managers at firm B gain $\$220.86 - \$134.78 = \$86.08$ by moving first. Managers at firm A gain the most from going first. If this were a question of acquiring the patent rights from an inventor, managers at firm A could outbid managers at firm B for the patent (because they can afford to bid up to \$89.88, whereas managers at firm B can afford to bid only \$86.08). In an ascending auction where all bids are public and the auction ends with the last bid unchallenged, we would expect managers at firm A to acquire the patent and pay a little more than \$86.08 for it (because they must only slightly outbid rivals to win the patent).

Oligopoly and competition among the few are so prevalent that these conditions should be studied carefully. For the first time in our analysis of firm behavior, *a manager's optimal strategy depends on what she thinks her adversary does*; that is, in equation (11.9) the marginal revenue depends not only on what the manager

does but also on what the manager's rival does. Hence the price a manager receives for her product depends on both her and her adversary's decisions. This is a chain of reciprocal decisions; that is, my actions depend on your actions, which in turn depend on my actions. Such interdependence of business decisions is typical of most of the economy. This mutual dependence is the basis of the strategy discussed in the game theory chapter (Chapter 12).

DUOPOLISTS AND PRICE COMPETITION WITH DIFFERENTIATED PRODUCTS

Is price competition always a lose–lose situation? We have shown that it is if there is no differentiation between market products. But what if managers can differentiate their products? Is price competition profitable? Let's view two competitors who produce *differentiated but highly substitutable products*. To keep the analysis simple, we'll assume the products have zero marginal cost. The demand for firm 1's product is expressed as

$$Q_1 = 100 - 3P_1 + 2P_2$$

where Q_1 is the quantity of firm 1's product demanded when managers price their product at P_1 per unit and managers at firm 2 price their product at P_2 per unit. Note that as managers at firm 2 price their product higher, the quantity demanded of firm 1's product increases as buyers switch. This shows that differentiation can only mitigate price competition; it is difficult to erase it as a purchase attribute. So again, the demand for firm 1's product depends not only on what managers control (their price) but also on what their rival charges (though they can influence the choices of rivals).

Analogously, the demand for firm 2's product is

$$Q_2 = 100 - 3P_2 + 2P_1$$

Managers at firm 1 want to maximize profit. In this case, this means maximizing the firm's total revenue because unit costs are 0.

As in Cournot, if managers get in a price war, they will compete prices down to marginal costs, and profit will be 0. Again, the price war is lose–lose. But is there a Cournot analog where they compete on price but don't commit economic suicide? There is, and it is called the *Bertrand model*. Firm 1's total revenue is

$$\begin{aligned} TR_1 &= P_1 * Q_1 = P_1 (100 - 3P_1 + 2P_2) = 100P_1 - 3P_1^2 + 2P_1P_2 \\ &= TR_{11} + TR_{12} \end{aligned}$$

where $TR_{11} = 100P_1 - 3P_1^2$ and $TR_{12} = 2P_1P_2$.

To maximize total revenue, we form $\Delta TR_1 / \Delta P_1 = (\Delta TR_{11} / \Delta P_1) + (\Delta TR_{12} / \Delta P_1)$ and set it equal to 0. $\Delta TR_{11} / \Delta P_1$ has the same form as when we did this earlier with respect to ΔQ —that is, $100 - 6P_1$ (same intercept of

100 and double the slope of -3). In the case of $\Delta TR_{12}/\Delta P_1$, call P_1 the initial price and P_1' the new price, so that $\Delta P_1 = P_1' - P_1$. Thus $\Delta TR_{12} = 2P_1'P_2 - 2P_1P_2 = 2P_2(P_1' - P_1) = 2P_2\Delta P_1$, and $\Delta TR_{12}/\Delta P_1 = 2P_2$. So the condition to maximize revenue (profit) for firm 1 will be

$$\Delta TR_1/\Delta P_1 = (100 - 6P_1) + 2P_2 = 0$$

or

$$6P_1 = 100 + 2P_2$$

or

$$P_1 = (50/3) + (1/3)P_2 \quad (11.12)$$

QUANT OPTION

The revenue [profit]-maximizing condition is

$$\partial TR_1/\partial P_1 = 100 - 6P_1 + 2P_2 = 0$$

In an analogous fashion, the profit-maximizing price for managers at firm 2 is

$$P_2 = (50/3) + (1/3)P_1 \quad (11.13)$$

Equations (11.12) and (11.13) give us two equations and two unknowns. Substituting (11.13) into (11.12) and solving yields

$$\begin{aligned} P_1 &= (50/3) + (1/3) \times [(50/3) + (1/3)P_1] = (150/9) + (50/9) + (1/9)P_1 \\ &= (200/9) + (1/9)P_1 \end{aligned}$$

or

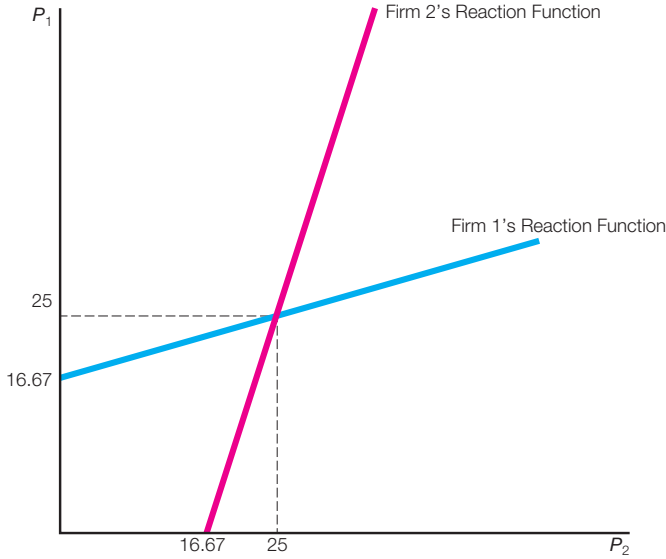
$$(8/9)P_1 = 200/9$$

Thus the optimal price for firm 1 managers to charge is $P_1 = 25$. Substituting $P_1 = 25$ in equation (11.13) gives $P_2 = (50/3) + (1/3)25 = 75/3 = 25$. Managers at both firms charge the same price because their differentiation efforts create similar impacts on the demand of others (remember their demand curves with the $+2$ in front of their rival's price). Substituting $P_1 = P_2 = 25$ into firm 1's demand curve gives $Q_1 = 100 - 3 \times 25 + 2 \times 25 = 100 - 75 + 50 = 75$ and analogously, $Q_2 = 75$. Managers at firm 1 earn total revenue (profit) of $TR_1 = P_1 \times Q_1 = 25 \times 75 = 1,875$ and analogously, $TR_2 = 1,875$. That's much nicer than the profit of 0 with the price war. The optimal solution is shown in Figure 11.5.

FIGURE 11.5

Bertrand Reaction Functions and Equilibrium for Firms 1 and 2

A Bertrand equilibrium occurs where the two firms' reaction functions intersect. This is the only price combination at which both firms' expectations of how the other firm will price are consistent with their own expectations of their own optimal price. In this case, both firms will price at 25.



What if the managers at the two firms showed more cooperative behavior and either colluded or merged? Then the combined entity's total revenue would be

$$\begin{aligned}
 TR &= TR_1 + TR_2 = P_1 \times (100 - 3P_1 + 2P_2) + P_2 \times (100 - 3P_2 + 2P_1) \\
 &= 100P_1 - 3P_1^2 + 2P_1P_2 + 100P_2 - 3P_2^2 + 2P_1P_2 \\
 &= 100P_1 - 3P_1^2 + 100P_2 - 3P_2^2 + 4P_1P_2 \\
 &= TR_{11} + TR_{22} + TR_{12}
 \end{aligned}$$

where

$$\begin{aligned}
 TR_{11} &= 100P_1 - 3P_1^2 \\
 TR_{22} &= 100P_2 - 3P_2^2 \\
 TR_{12} &= 4P_1P_2
 \end{aligned}$$

To maximize total revenue (profit), managers will set $\Delta TR / \Delta P_1 = 0$ and $\Delta TR / \Delta P_2 = 0$.

Analogous to the procedure for deriving equations (11.12) and (11.13), we set

$$\Delta(TR_{11} + TR_{12}) / \Delta P_1 = \Delta TR / \Delta P_1 = 100 - 6P_1 + 4P_2 = 0$$

or

$$6P_1 = 100 + 4P_2$$

or

$$P_1 = (50/3) + (2/3)P_2 \quad (11.14)$$

By setting $\Delta(TR_{22} + TR_{12})/\Delta P_2 = \Delta TR/\Delta P_2 = 0$, the optimal price for P_2 is

$$P_2 = (50/3) + (2/3)P_1 \quad (11.15)$$

QUANT OPTION

To maximize revenue [profit], the firm will set

$$\partial TR/\partial P_1 = 100 - 6P_1 + 4P_2 = 0$$

and

$$\partial TR/\partial P_2 = 100 - 6P_2 + 4P_1 = 0$$

With two equations and two unknowns, we solve by substituting (11.15) into (11.14)

$$\begin{aligned} P_1 &= (50/3) + (2/3)[(50/3) + (2/3)P_1] = (150/9) + (100/9) + (4/9)P_1 \\ &= (250/9) + (4/9)P_1 \end{aligned}$$

or

$$(5/9)P_1 = 250/9$$

This results in $P_1 = 50$ and $P_2 = 50$. Substituting $P_1 = P_2 = 50$ into firm 1's demand gives $Q_1 = 100 - 3 \times 50 + 2 \times 50 = 100 - 150 + 100 = 50$ and analogously $Q_2 = 50$. Managers at firm 1 earn total revenue (profit) of $TR_1 = P_1 \times Q_1 = 50 \times 50 = 2,500$. Analogously, $TR_2 = 2,500$. Clearly collusion yields better profit, but it is also generally illegal. Managers who choose to compete on prices yield $(1,875/2,500) = 75\%$ of the collusion profit and avoid the price war (as well as jail). Note that turning this around as we did with Cournot, the entry of another competitor reduces prices (in this example by 50%). This is why managers need to think carefully about the effects of competition on variable profit. Competition can allow consumers to keep more consumer surplus and reduce producer surplus. We will further discuss strategic pricing when competition exists among a few firms in Chapter 12.

THE STICKY PRICING OF MANAGERS

The Cournot model explains why price may be “sticky”; that is, managers evolve toward the optimum and stay there. Even in markets with homogeneous products, managers show little incentive to deviate. This is especially true in markets where cost and demand have been stable or easily anticipated and managers have competed for several years. Another behavioral model explains why prices can be sticky even when products are somewhat differentiated.

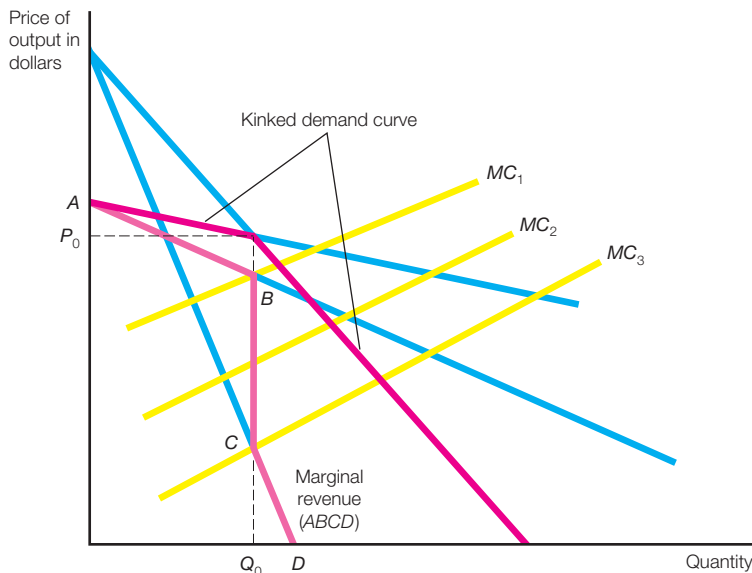
Consider managers facing a limited number of competitors. They currently price at P_0 and produce Q_0 . Should managers increase price, demand will be quite elastic (but not perfectly elastic because with a differentiated product, rival products are not perfect substitutes). Some customers will buy elsewhere when price increases, but other customers have a higher value for the product.

On the other hand, should managers drop their price, they could assume that demand will become less elastic because rivals will also reduce prices to protect their sales. Although lowering the price, *if* no other firm followed suit, might

FIGURE 11.6

The Situation of the Kinked Demand Curve

The demand curve kinks at (Q_0, P_0) , with the curve being relatively elastic above the kink and relatively less elastic below the kink. The marginal revenue curve is discontinuous at Q_0 (gap BC), and the marginal cost curves intersect the marginal revenue in the gap, leading to marginal revenue equal to marginal cost and yielding an optimal price of P_0 and an optimal quantity of Q_0 , despite major shifts in the marginal cost curve.



increase sales, when rivals follow a price cut, margins decrease and the increase in sales may not make up the difference.

Hence managers face a kinked demand curve at point (Q_0, P_0) , with demand being gently sloped above it and steeply sloped below it. This pattern yields a discontinuous marginal revenue curve (see gap BC in Figure 11.6) at Q_0 (recall that for linear demand curves, the marginal revenue curves have the same dollar axis intercept but fall twice as fast).

Therefore, the marginal cost curves MC_1 and MC_3 (and anything in between, such as MC_2) yield the same price P_0 and quantity Q_0 for the profit-maximizing profit when they intersect the marginal revenue ($ABCD$) in the discontinuity. Thus costs can shift around quite a bit without changing the profit-maximizing price (making it sticky).

SUMMARY

1. Oligopolistic markets are characterized by a small number of firms with a great deal of interdependence, actual and perceived. A good example of an oligopoly is the U.S. oil industry, in which a few firms account for the bulk of the industry's capacity.

2. There is no single model of oligopolistic behavior; behavior depends on the circumstances and expectations. Conditions in oligopolistic industries tend to promote cooperative behavior among rivals, including collusion. Cooperation is easier to attain because the number of firms is small, and managers recognize their interdependence with rivals. The advantages to be derived by managers from collusion seem obvious: increased profit, decreased uncertainty, and a better opportunity to control the entry of new firms. However, collusive arrangements are often hard to maintain because once a collusive agreement is made, any member can increase its profit by cheating on the agreement. Also, cartels may find it difficult to identify a course of action that is agreeable to all members of the industry.

3. Another model of oligopolistic behavior is price leadership. The price leader is the dominant firm in the market. We showed how, under these circumstances, managers of the price leader will set prices to maximize profit. This model also explains pricing in an imperfect cartel (where not all producers are members).

4. When competition among a few firms exists, managers may engage in price competition. This is usually a lose-lose situation because prices are competed down to marginal costs, with severe impact on profit. Cournot competition (competition on quantity or capacity) is a strategy to capture a significant percentage of the high cartel profit and avoid the negative impact on profit of price competition. Cournot behavior exists when managers move simultaneously and engage in what-if moves; for example, what would be my profit-maximizing output response given an output by my rivals? By logically tracing all profit-maximizing responses to rivals' output choices, then putting yourself in their posi-

tion and doing the same analysis for them, behavior can become consistent across all market players. The key to competition among the few is interdependence, where your optimal output is a function not only of what you wish to do but also what your rivals wish to do.

5. If managers at one firm can act before the managers of another firm, Stackelberg behavior can explain the optimal (profit-maximizing) strategy for the first mover and all subsequent movers in this sequential game. In general, first movers see their profit improve relative to the simultaneous-move Cournot situation, and subsequent movers see their profit decrease relative to the Cournot situation. If managers of low-cost firms move first, they earn higher profits than managers of high-cost first movers. If managers can purchase the right to move first, the value they place on this purchase is their profit if they move first minus their profit if they follow.

6. Managers with differentiated products may compete on price and not compete price down to marginal cost. If a limited number of such firms compete, they can determine their and their rivals' reaction functions. Their profit-maximizing price will depend on the prices of rivals as well as their own price. Managers can deduce the optimal price they should charge because they can deduce the optimal prices of rivals. The optimal prices are considerably lower than the price the managers would charge if they formed a cartel and considerably higher than if the managers engaged in a price war.

7. Prices may be sticky (that is, tend to be stable) in oligopolies with differentiated products. This occurs because the demand curve kinks at the current price. The curve is very elastic above the current price because rivals do not follow the price increases of others; whereas it is much less elastic below the current price because rivals are likely to meet any price decreases to protect their sales. The kink in the demand curve leads to a discontinuity in the firm's marginal revenue curve. Therefore, the marginal cost can shift upward or downward considerably but still meet the marginal revenue equals marginal cost condition for profit maximization in the marginal revenue discontinuity—and thus not change the profit-maximizing price and quantity.

PROBLEMS

1. The Bergen Company and the Gutenberg Company are the only two firms that produce and sell a particular kind of machinery. The demand curve for their product is

$$P = 580 - 3Q$$

where P is the price (in dollars) of the product, and Q is the total amount demanded. The total cost function of the Bergen Company is

$$TC_B = 410Q_B$$

where TC_B is its total cost (in dollars) and Q_B is its output. The total cost function of the Gutenberg Company is

$$TC_G = 460Q_G$$

where TC_G is its total cost (in dollars) and Q_G is its output.

- a. If these two firms collude and they want to maximize their combined profit, how much will the Bergen Company produce?
 - b. How much will the Gutenberg Company produce?
 - c. Will the Gutenberg Company agree to such an arrangement? Why or why not?
2. The can industry is composed of two firms. Suppose that the demand curve for cans is

$$P = 100 - Q$$

where P is the price (in cents) of a can and Q is the quantity demanded (in millions per month) of cans. Suppose the total cost function of each firm is

$$TC = 2 + 15q$$

where TC is total cost (in tens of thousands of dollars) per month and q is the quantity produced (in millions) per month by the firm.

- a. What are the price and output if managers set price equal to marginal cost?
 - b. What are the profit-maximizing price and output if the managers collude and act like a monopolist?
 - c. Do the managers make a higher combined profit if they collude than if they set price equal to marginal cost? If so, how much higher is their combined profit?
3. An oligopolistic industry selling a particular type of machine tool is composed of two firms. Managers at the two firms set the same price and share the total market equally. The demand curve confronting each firm (assuming that the other firm sets the same price) follows, as well as each firm's total cost function.

Price (Thousands of Dollars)	Quantity Demanded per Day	Daily Output	Total Cost (Thousands of Dollars)
10	5	5	45
9	6	6	47
8	7	7	50
7	8	8	55
6	9	9	65

- a. Assuming that each manager is correct in believing that managers at the other firm will charge the same price as they do, what price should each charge?
 - b. Under the assumptions in part (a), what daily output rate should managers at each firm set?
4. James Pizzo is president of a firm that is the industry price leader; that is, it sets the price and the other firms sell all they want at that price. In other words, the other firms act as perfect competitors. The demand curve for this industry's product is $P = 300 - Q$, where P is the price of the product and Q is the total quantity demanded. The total amount supplied by the other firms is equal to Q_r , where $Q_r = 49P$. (P is measured in dollars per barrel; Q , Q_r , and Q_b are measured in millions of barrels per week.)
 - a. If Pizzo's firm's marginal cost curve is $2.96Q_b$, where Q_b is the output of his firm, at what output level should he operate to maximize profit?
 - b. What price should he charge?
 - c. How much does the industry as a whole produce at this price?
 - d. Is Pizzo's firm the dominant firm in the industry?
 5. The International Air Transport Association (IATA) has been composed of 108 U.S. and European airlines that fly transatlantic routes. For many years, IATA acted as a cartel: It fixed and enforced uniform prices.
 - a. If IATA wanted to maximize the total profit of all member airlines, what uniform price would it charge?
 - b. How would the total amount of traffic be allocated among the member airlines?
 - c. Would IATA set price equal to marginal cost? Why or why not?
 6. In late 1991 two firms, Delta Airlines and the Trump Shuttle, provided air shuttle service between New York and Boston or Washington. The one-way price charged by both firms was \$142 on weekdays and \$92 on weekends, with lower off-peak advance purchase fares. In September 1991 Delta increased the per-trip shuttle mileage given to members of the Delta frequent-flier program from 1,000 to 2,000 miles, even though actual mileage from New York to either Boston or Washington is about 200 miles. Moreover, Delta also offered an extra 1,000 miles to frequent fliers who made a round-trip on the same day, raising a possible day's total to 5,000 miles. Almost simultaneously, Trump changed the frequent-flier mileage it gave shuttle passengers. (It participated in the One Pass frequent-flier program with Continental Airlines and some foreign carriers.) What sorts of changes do you think Trump made? Why?
 7. Two firms, the Alliance Company and the Bangor Corporation, produce vision systems. The demand curve for vision systems is

$$P = 200,000 - 6(Q_1 + Q_2)$$

where P is the price (in dollars) of a vision system, Q_1 is the number of vision systems produced and sold per month by Alliance, and Q_2 is the number of

vision systems produced and sold per month by Bangor. Alliance's total cost (in dollars) is

$$TC_1 = 8,000Q_1$$

Bangor's total cost (in dollars) is

$$TC_2 = 12,000Q_2$$

- a. If managers at these two firms set their own output levels to maximize profit, assuming that managers at the other firm hold constant their output, what is the equilibrium price?
 - b. What is the output of each firm?
 - c. How much profit do managers at each firm earn?
8. In Britain price competition among bookshops has been suppressed for over 100 years by the Net Book Agreement (of 1900), which was aimed at preventing price wars. However, in October 1991 Waterstone and Company began cutting book prices at its 85 British shops. According to Richard Barker, Waterstone's operations director, the decision to reduce the price of about 40 titles by about 25% was due to price cuts by Dillons, Waterstone's principal rival.
- a. According to the president of Britain's Publishers Association, the price-cutting was "an enormous pity" that will "damage many booksellers who operate on very slim margins."⁷ Does this mean that price-cutting of this sort is contrary to the public interest?
 - b. Why would Dillons want to cut prices? Under what circumstances would this be a good strategy? Under what circumstances would it be a mistake?
9. In the 1960s Procter & Gamble recognized that disposable diapers could be made a mass-market product and developed techniques to produce diapers at high speed and correspondingly low cost. The result was that it dominated the market. According to Harvard's Michael Porter, who made a careful study of this industry, the following were some ways in which Procter & Gamble might have signaled other firms to deter entry.⁸

7. Suzanne Cassidy, "British Book Shops in Price Skirmishes," *New York Times*, October 7, 1991, at www.nytimes.com/1991/10/07/business/the-media-business-british-book-shops-in-price-skirmishes.html.

8. M. Porter, "Strategic Interaction: Some Lessons from Industry Histories for Theory and Antitrust Policy," in S. Salop, ed., *Strategy, Predation, and Antitrust Analysis* (Washington, DC: Federal Trade Commission, 1981); *New York Times*, April 15, 1993, and March 25, 1995; and *BusinessWeek*, April 26, 1993, and September 19, 1994.

Tactic	Cost to Procter & Gamble	Cost to an Entrant
1. Signal a commitment to defend position in diapers through public statements, comments to retailers, etc.	None	Raises expected cost of entry by increasing probability and extent of retaliation.

2. File a patent suit.	Legal fees	Incurs legal fees plus the probability that P & G wins the suit with subsequent cost to the competitor.
3. Announce planned capacity expansion.	None	Raises expected risk of price-cutting and the probability of P & G's retaliation to entry.
4. Announce a new generation of diapers to be introduced in the future.	None	Raises the expected cost of entry by forcing entrant to bear possible product development and changeover costs contingent on the ultimate configuration of the new generation.

- a. In considering these possible tactics, why should managers at Procter & Gamble be concerned about their costs?
 - b. Why should managers be concerned with the costs to an entrant?
 - c. By the 1990s Procter & Gamble had to compete with high-quality, private-label diapers (as well as with Kimberly-Clark, which successfully entered the market in the 1970s). In March 1993 its Pampers brand had about 30% of the market, and its Luvs brand had about 10%. The price of Luvs and Pampers exceeded that of discount brands by over 30%. Should Procter & Gamble have cut its prices?
 - d. In 1993 Procter & Gamble sued Paragon Trade Brands, a private-label producer, alleging infringement of two patents. Are lawsuits of this kind part of the process of oligopolistic rivalry and struggle?
10. Under which circumstances do managers find it profitable to increase the quality of their products? Do the benefits always exceed the costs? Why or why not?
 11. The West Chester Corporation believes that the demand curve for its product is

$$P = 28 - 0.14Q$$

where P is price (in dollars) and Q is output (in thousands of units). The firm's board of directors, after a lengthy meeting, concludes that the firm should

attempt, at least for a while, to increase its total revenue, even if this means lower profit.

- a. Why might managers adopt such a policy?
 - b. What price should managers set if they want to maximize total revenue?
 - c. If the firm's marginal cost equals \$14, do managers produce a larger or smaller output than they would to maximize profit? How much larger or smaller?
12. Steve Win has purchased land from the city of Atlantic City in the Marina section. There are stories of a new casino building boom in Atlantic City (MGeeM is also talking about entering, and Gump is opening his fourth casino). Some talk is circulating that Win will subdivide his new land purchase and perhaps three casinos will be built on the site.

Suppose Win subdivides his land into two parcels. He builds on one site and sells the other to another gambling entrepreneur. Win estimates that the demand for gambling in the Marina area of Atlantic City (*after* accounting for the presence of two existing casinos in the Marina and adjusting for the rest of the casinos in Atlantic City) is

$$P = 750 - 5Q$$

where P is the price associated with gambling and Q is the quantity of gambling (think of P as the average amount that a typical patron will net the casino, an amount paid for the entertainment of gambling, and Q as the number of gamblers).

Win, of course, does not sell the other parcel until his casino is built (or is significantly far along); thus he has a first-mover advantage.

Win's total cost TC_W of producing gambling is

$$TC_W = 20 + 40Q_W + 15.5Q_W^2$$

where Q_W is the number of gamblers in Win's casino and the total cost (TC_R) of producing gambling for Win's rival is

$$TC_R = 10 + 50Q_R + 20Q_R^2$$

where Q_R is the number of gamblers in the rival's casino and

$$Q_W + Q_R = Q$$

Would Atlantic City have done better to sell the land as two separate parcels rather than as a single parcel to Win (given that Win was going to subdivide, Win and his rival could not collude, and Win did not have the ability to produce as a monopolist)? You may assume that Win and his rival could have been Cournot duopolists. If Atlantic City could do better, show why and by how much. Carry all calculations to the thousandths decimal point.

EXCEL EXERCISE: DOMINANT FIRM PRICE LEADER

Suppose that the market demand for a product was expressed as $Q_M = 15,000 - 3,000P$. The follower firms have a collective marginal cost of $MC_F = 3 + 0.001Q_F$. Because the follower firms will take the price the leader sets as given, follower firms will maximize profits when

$$P = MC_F = 3 + 0.001Q_F$$

or
$$0.001Q_F = -3 + P$$

or
$$Q_F = -3,000 + 1,000P$$

The leader's residual demand curve is $Q_L = Q_M - Q_F$ or

$$\begin{aligned} Q_L &= 15,000 - 3,000P - (-3,000 + 1,000P) \\ &= 15,000 - 3,000P + 3,000 - 1,000P = 18,000 - 4,000P \end{aligned}$$

or
$$4,000P = 18,000 - Q_L$$

$$P = 4.5 - 0.00025Q_L$$

The leader's marginal revenue is the same intercept and double the slope (as described in the text) and so

$$MR_L = 4.5 - 0.0005Q_L$$

Suppose the leader's marginal cost (and average variable cost) is $MC_L = 2.5 = AVC_L$.

Suppose the leader's fixed cost is $FC_L = 1,000$.

To maximize profits, the leader will set $MR_L = MC_L$ or

$$MR_L = 4.5 - 0.0005Q_L = 2.5 = MC_L$$

or
$$0.0005Q_L = 2$$

or
$$Q_L = 4,000$$

Substituting $Q_L = 4,000$ into the leader's demand curve gives

$$P = 4.5 - 0.00025 \cdot 4,000 = 4.5 - 1 = 3.5$$

Substituting $P = 3.5$ into the follower's supply curve gives

$$Q_F = -3,000 + 1,000 \cdot 3.5 = -3,000 + 3,500 = 500$$

The leader's total revenue = $TR_L = P \cdot Q_L = 3.5 \cdot 4,000 = 14,000$.

The leader's variable cost = $VC_L = AVC_L \cdot Q_L = 2.5 \cdot 4,000 = 10,000$.

The leader's profit is $\pi_L = TR_L - VC_L - FC_L = 14,000 - 10,000 - 1,000 = 3,000$.

Suppose that we didn't know the above formulas and the calculus rules that lead us to the above optimal behavior by the dominant firm. Instead, we have the following data about estimated market demand and estimated follower supply from the leader firm's marketing group. The leader firm's cost folks tell us that there is a constant marginal cost (= average variable cost) of 2.5, and a fixed cost of 1,000. Can we use a spreadsheet to answer this question? By now, you know the answer is yes.

Price	Market Demand	Follower Supply
5	0	2,000
4.75	750	1,750
4.5	1,500	1,500
4.25	2,250	1,250
4	3,000	1,000
3.75	3,750	750
3.5	4,500	500
3.25	5,250	250
3	6,000	0
2.75	6,750	0
2.5	7,500	0
2.25	8,250	0
2	9,000	0
1.75	9,750	0
1.5	10,500	0
1.25	11,250	0
1	12,000	0
0.75	12,750	0
0.5	13,500	0
0.25	14,250	0
0	15,000	0

Open your spreadsheet. Enter 5 in cell A1, enter 4.75 in cell A2, and so on, until you've entered 0 in cell A21.

Enter 0 in cell B1, 750 in cell B2, and so on, until you've entered 15000 in cell B21.

Columns A and B give you the market demand curve.

Enter 2000 in cell C1, 1,750 in cell C2, and so on, until you enter 0 in cell C9. Then click on the lower right-hand corner of cell C9 and drag your mouse down to cell C21. Columns A and C give you the followers' supply curve.

We know that the leader will allow the followers to supply as much as they want *at the price the leader sets*. As can be seen from the table above, at a price of 4.50, the followers will supply exactly what the market demands. At the two prices above 4.50 (4.75 and 5), the followers will supply more than the market

demands. Thus, at those prices (4.50, 4.75, and 5), there will be no demand for the leader. However, at prices below 4.5, the followers will provide less than the market demands, and thus the residual will be left for the leader. So here's what we do:

Enter 0 in cell D1, 0 in cell D2, =B3–C3 in cell D3, and so on via click and drag, until cell D21. Column D gives you the leader's quantity demanded at the corresponding price in column A.

Enter =A1*D1 in cell E1, and so on via click and drag, until cell E21. Column E gives the leader's total revenue.

Enter =2.5*D1 in cell F1, = 2.5*D2 in cell F2, and so on via click and drag, until cell F21. Column F is the leader's variable cost.

Enter 1,000 in cell G1, and so on via click and drag, until cell G21. Column G is the leader's fixed cost.

Enter =F1+G1 in cell H1, and so on via click and drag, until cell H21. Column H is the leader's total cost.

Enter =E1–H1 in cell I1, and so on via click and drag, until cell I21. Column I is the leader's profit. Then search column I for the highest value (or put =Max (I1:I21) in cell I22 and let the spreadsheet find the highest value for you). It'll be 3,000 in cell I7. This is the same result obtained by using the calculus-derived formulas. Going leftward on row 7, you will see the leader's output at 4,000, the follower's output at 500, and the optimal market price at 3.5 (as we found above with the calculus-based formulas).

Suppose that some members of the leader's management team had been pressuring the firm to eliminate the follower firms. This could be done by pricing at 3 (note that the follower's supply is zero if the price is 3). This would eliminate competition, increase quantity sold (to 6,000 from 4,000—a 50% increase), and still allow a positive profit margin of 0.5 per unit (=3.0–2.5). However, this would be a drop of 50% from the current profit margin of 1 (=3.5–2.5). Should the management of the leader firm drive out the competitors by pricing at 3? The spreadsheet tells us no. By pricing at 3, the leader's profit falls by a third (from 3,000 to 2,000). The increase in quantity sold is not enough to offset the decrease in profit margin.

It's often the case that the price leader will tolerate fringe players in the market. It's not because the leader is just a nice entity—it's because it's more profitable for the leader to have the fringe there.

EXCEL EXERCISE: COURNOT

Suppose that the market demand for the product was $P = 14 - Q$ or $P = 14 - Q_1 - Q_2$. Each of the two firms has marginal cost (MC) of 2 and average variable cost (AVC) of 2. They also each have fixed cost (FC) of 10.

Firm 1's total revenue is

$$TR_1 = P \cdot Q_1 = (14 - Q_1 - Q_2) \cdot Q_1 = 14Q_1 - Q_1^2 - Q_1Q_2$$

Firm 1's marginal revenue is

$$\partial TR_1 / \partial Q_1 = MR_1 = 14 - 2Q_1 - Q_2$$

To maximize profit, firm 1 will set marginal revenue equal to marginal cost, that is,

$$MR_1 = 14 - 2Q_1 - Q_2 = 2 = MC_1$$

or $2Q_1 = 12 - Q_2$

or $Q_1 = 6 - 0.5Q_2$

This is firm 1's reaction function.

In analogous fashion, firm 2's reaction function is $Q_2 = 6 - 0.5Q_1$.

Substituting firm 2's reaction function into firm 1's reaction function gives

$$Q_1 = 6 - 0.5(6 - 0.5Q_1) = 6 - 3 + 0.25Q_1 = 3 + 0.25Q_1$$

or $0.75Q_1 = 3$

or $Q_1 = 4$

Substituting $Q_1 = 4$ into firm 2's reaction function gives

$$Q_2 = 6 - 0.5 \cdot 4 = 6 - 2 = 4$$

Then $Q = Q_1 + Q_2 = 4 + 4 = 8$.

Substituting $Q = 8$ into the market demand function gives

$$P = 14 - Q = 14 - 8 = 6$$

Thus, each firm makes the following profits.

$$\text{Total revenue} = P \cdot Q = 6 \cdot 4 = 24$$

$$\text{Variable Cost} = AVC \cdot Q = 2 \cdot 4 = 8$$

$$\text{Fixed Cost} = FC = 10$$

$$\text{Profit } \pi = TR - VC - FC = 24 - 8 - 10 = 6$$

This is the Cournot solution if you know the calculus. But what if you don't know the calculus, and you only have the information that each firm has a marginal cost of 2, an average variable cost of 2, a fixed cost of 10, and that the market demand curve is as below.

Price	Quantity
14	0
13	1
12	2
11	3
10	4
9	5

8	6
7	7
6	8
5	9
4	10
3	11
2	12
1	13
0	14

Let's suppose that you are firm 1. You start by asking yourself "what if" questions. This is scenario planning. Early in the game, you don't know what your adversary is going to do. But you want to consider all possible actions your adversary might take and what profit-maximizing action you would take in response. We'll do that now with a spreadsheet.

Suppose your adversary abdicates the market to you. You then face the market demand curve and act as a profit-maximizing monopolist.

Open your spreadsheet. Enter 14 in cell A1, 13 in cell A2, and so on, until you've entered 0 in cell A15.

Enter 0 in cell B1, 1 in cell B2, and so on, until you've entered 14 in cell B15.

You now have the firm's demand curve in Columns A and B.

Enter $=A1*B1$ in cell C1. Then click on the lower right-hand corner of cell C1 with your mouse and drag down to cell C15. The result will give you a column of the firm's total revenue.

Enter $=2*B1$ in cell D1 and then click and drag to cell D15. Column D is the firm's variable cost.

Enter 10 in cell E1 and then click and drag until cell E15. Column E is the firm's fixed cost.

Enter $=D1+E1$ in cell F1 and then click and drag until cell F15. Column F is the firm's total cost.

Enter $=C1-F1$ in column G1 and then click and drag down to cell G15. Column G is the firm's profit. Then search column G for the highest number (or enter $=\text{Max}(G1:G15)$ in cell G16 and let the spreadsheet find the highest number). This will be 26 in cell G7. Reading leftward in the seventh row, you will see that you will produce 6 units in cell B7. Thus, you have answered one "what-if" question: If my adversary produces 0, I will produce 6.

Let's do another what-if scenario. What if your adversary produces 2? Then what's left for you is the market demand less 2. For instance, at market price 12, 2 goods are demanded. But if your adversary produces 2, there is nothing for you. At market price 11, there are 3 goods demanded. If your adversary produces 2, there is 1 left for you. We continue with this until we get to market price 0. At that

price 14 are demanded, but since your adversary is producing 2, there are 12 available to you. So to get your demand curve *given that your adversary is producing two*, we take the market demand curve (Columns A and B), and at any given price, we subtract 2 from the amount in column B. Since that would entail negative quantities for prices 14 and 13 (which makes no sense), we start the process at price 12.

Enter 0 in cell H3, 1 in cell H4, and so on, until you enter a 12 in cell H15. Column A and Column H yield your “residual demand curve,” that is, what’s left for you of the market demand after your adversary has taken 2.

Enter $=A3*H3$ in cell I3 and then click and drag to cell I15. Column I is your total revenue at any price when your adversary produces 2.

Enter $=10+2*H3$ in cell J3 and then click and drag to cell J15. Column J is your total cost.

Enter $=I3-J3$ in cell K3 and then click and drag to cell K15. Column K is your profit if your adversary produces 2. Search column K for the largest number (or let the spreadsheet do it for you). It will be 15 in cell K8. Looking leftward in the eighth row, you will see that your optimal output (in cell H8) is 5. Thus, if your adversary produces 2, you would produce 5 to maximize your profit.

We continue with another what-if scenario. What if your adversary produces 4? Then what’s left for you is the market demand less 4. For instance, at market price 10, 4 goods are demanded. But if your adversary produces 4, there is nothing for you. At market price 9, there are 5 goods demanded. If your adversary produces 4, there is 1 left for you. We continue with this until we get to market price 0. At that price, 14 are demanded, but since your adversary is producing 4, there are 10 available to you. So to get your demand curve given that your adversary is producing 4, we take the market demand curve (Columns A and B), and at any given price, we subtract 4 from the amount in column B. Since that would entail negative quantities for prices between 14 and 11 (which makes no sense), we start the process at price 10.

Enter 0 in cell L5, 1 in cell L6, and so on, until you enter a 10 in cell L15. Column A and Column L yield your “residual demand curve,” that is, what’s left for you of the market demand after your adversary has taken 4.

Enter $=A5*L5$ in cell M5 and then click and drag to cell M15. Column M is your total revenue at any price when your adversary produces 4.

Enter $=10+2*L5$ in cell N5 and then click and drag to cell N15. Column N is your total cost.

Enter $=M5-N5$ in cell O5 and then click and drag to cell O15. Column O is your profit if your adversary produces 4. Search column O for the largest number (or let the spreadsheet do it for you). It will be 6 in cell O9. Looking leftward in the ninth row, you will see that your optimal output (in cell L9) is 4. Thus, if your adversary produces 4, you would produce 4 to maximize your profit.

What if your adversary produces 6? Then what's left for you is the market demand less 6. For instance, at market price 8, 6 goods are demanded. But if your adversary produces 6, there is nothing for you. At market price 7, there are 7 goods demanded. If your adversary produces 6, there is 1 left for you. We continue with this until we get to market price 0. At that price 14 are demanded, but since your adversary is producing 6, there are 8 available to you. So to obtain your demand curve based on your adversary producing 6, we take the market demand curve (Columns A and B), and at any given price, we subtract 6 from the amount in column B. Since that would entail negative quantities for prices between 14 and 9 (which makes no sense), we start the process at price 8.

Enter 0 in cell P7, 1 in cell P8, and so on, until you enter a 8 in cell P15. Column A and Column P yield your "residual demand curve," that is, what's left for you of the market demand after your adversary has taken 6.

Enter $=A7*P7$ in cell Q7 and then click and drag to cell Q15. Column Q is your total revenue at any price when your adversary produces 6.

Enter $=10+2*P7$ in cell R7 and then click and drag to cell R15. Column R is your total cost.

Enter $=Q7-R7$ in cell S7 and then click and drag to cell S15. Column S is your profit if your adversary produces 6. Search column S for the largest number (or let the spreadsheet do it for you). It will be -1 in cell S10. Looking leftward in the tenth row, you will see that your optimal output (in cell P10) is 3. Thus, if your adversary produces 6, you would produce 3 to maximize your profit. You might wonder why you would produce at all given that your profit is negative. Suppose that the 10 in fixed cost was sunk. Then your variable profit (cash flow under your control) is profit plus fixed cost, which is 9 ($=-1 + 10$). Thus, you would continue producing.

What if your adversary produces 8? Then what's left for you is the market demand less 8. For instance, at market price 6, 8 goods are demanded. But if your adversary produces 8, there is nothing for you. At market price 5, there are 9 goods demanded. If your adversary produces 8, there is 1 left for you. We continue with this until we get to market price 0. At that price, 14 are demanded, but since your adversary is producing 8, there are 6 available to you. So to get your demand curve given that your adversary is producing 8, we take the market demand curve (Columns A and B), and at any given price we subtract 8 from the amount in column B. Since that would entail negative quantities for prices between 14 and 7 (which makes no sense), we start the process at price 6.

Enter 0 in cell T9, 1 in cell T10, and so on, until you enter a 6 in cell T15. Column A and Column T yield your "residual demand curve," that is, what's left for you of the market demand after your adversary has taken 8.

Enter $=A9*T9$ in cell U9 and then click and drag to cell U15. Column U is your total revenue at any price when your adversary produces 8.

Enter $=10+2*T9$ in cell V9 and then click and drag to cell V15. Column V is your total cost.

Enter $=U9-V9$ in cell W9 and then click and drag to cell W15. Column W is your profit if your adversary produces 8. Search column W for the largest number (or let the spreadsheet do it for you). It will be -6 in cell W11. Looking leftward in the eleventh row shows you that your optimal output (in cell T11) is 2. Thus, if your adversary produces 8, you would produce 2 to maximize your profit. You might wonder why you would produce at all given that your profit is negative. Suppose that the 10 in fixed cost was sunk. Then your variable profit (cash flow under your control) is profit plus fixed cost, which is 4 ($=-6 + 10$). Thus, you would continue producing.

What if your adversary produces 10? Then what's left for you is the market demand less 10. For instance, at market price 4, 10 goods are demanded. But if your adversary produces 10, there is nothing for you. At market price 3, there are 11 goods demanded. If your adversary produces 10, there is 1 left for you. We continue with this until we get to market price 0. At that price, 14 are demanded, but since your adversary is producing 10, there are 4 available to you. So to get your demand curve given that your adversary is producing 10, we take the market demand curve (Columns A and B), and at any given price, we subtract 10 from the amount in column B. Since that would entail negative quantities for prices between 14 and 5 (which makes no sense), we start the process at price 4.

Enter 0 in cell X11, 1 in cell X12, and so on until you enter a 4 in cell X15. Column A and Column X yield your "residual demand curve," that is, what's left for you of the market demand after your adversary has taken 10.

Enter $=A11*X11$ in cell Y11 and then click and drag to cell Y15. Column Y is your total revenue at any price when your adversary produces 10.

Enter $=10+2*X11$ in cell Z11 and then click and drag to cell Z15. Column Z is your total cost.

Enter $=Y11-Z11$ in cell AA11 and then click and drag to cell AA15. Column AA is your profit if your adversary produces 10. Search column AA for the largest number (or let the spreadsheet do it for you). It will be -9 in cell AA12. Looking leftward in row 12 shows you that your optimal output (in cell X12) is 1. Thus, if your adversary produces 10, you would produce 1 to maximize your profit. You might wonder why you would produce at all given that your profit is negative. Suppose that the 10 in fixed cost was sunk. Then your variable profit (cash flow under your control) is profit plus fixed cost, which is 1 ($=-9 + 10$). Thus, you would continue producing.

What if your adversary produces 12? Then what's left for you is the market demand less 12. For instance, at market price 2, 12 goods are demanded. But if your adversary produces 12, there is nothing for you. At market price 1, there are 13 goods demanded. If your adversary produces 12, there is 1 left for you. We continue with this until we get to market price 0. At that price, 14 are demanded,

but since your adversary is producing 12, there are 2 available to you. So to get your demand curve given that your adversary is producing 12, we take the market demand curve (Columns A and B), and at any given price, we subtract 12 from the amount in column B. Since that would entail negative quantities for prices between 14 and 3 (which makes no sense), we start the process at price 2.

Enter 0 in cell AB13, 1 in cell AB14, and 2 in cell AB15. Columns A and AB yield your “residual demand curve,” that is, what’s left for you of the market demand after your adversary has taken 12.

Enter $=A13*AB13$ in cell AC13 and then click and drag to cell AC15. Column AC is your total revenue at any price when your adversary produces 12.

Enter $=10+2*AB13$ in cell AD13 and then click and drag to cell AD15. Column AD is your total cost.

Enter $=AC13-AD13$ in cell AE13 and then click and drag to cell AE15. Column AE is your profit if your adversary produces 12. Search column AE for the largest number (or let the spreadsheet do it for you). It will be -10 in cell AE13. Looking leftward in the thirteenth row, you will see that your optimal output (in cell AB13) is 0. Thus, if your adversary produces 12, you would produce 0 to maximize your profit. If you produce 0 when your adversary produces 12, you will produce 0 if your adversary produces more than 12.

Now let’s summarize what we’ve found by asking our “what-ifs.”

If My Adversary Produces (Q_2)	My Optimal Output is (Q_1)
0	6
2	5
4	4
6	3
8	2
10	1
12	0

From this table, we notice that every time my adversary produces two more, I want to produce one less. We can express this by an equation: $Q_1 = 6 - 0.5Q_2$, or we could plot the above points on a graph with Q_1 on the y axis and Q_2 on the x axis. The y intercept would be a 6 (where $Q_2 = 0$), and the x intercept would be at 12 (where $Q_1 = 0$). The slope is the change in the y axis over the change in the x axis (between any two points—the “rise over the run,” i.e., $-6/12 = -0.5$). Thus, the equation of the line we graphed is $Q_1 = 6 - 0.5Q_2$. Notice this is firm 1’s reaction function that we derived from the calculus at the beginning of the exercise. We computed it here without any calculus.

We could do the same analysis for my adversary, and we would come up with a comparable table.

If I Produce Q_1	My Adversary's Optimal Output is Q_2
0	6
2	5
4	4
6	3
8	2
10	1
12	0

As above, we can see that when my output goes up by two, my adversary's output goes down by one, or we could graph the points. Either way would lead to firm 2's reaction function of $Q_2 = 6 - 0.5Q_1$.

We can then observe that the only consistent output for each firm that is an equilibrium is for both firms to produce 4. Suppose I produce 6. Then my adversary would want to produce $Q_2 = 6 - 0.5*6 = 6 - 3 = 3$. But then based on his or her reaction function, I would want to produce $Q_1 = 6 - 0.5*3 = 6 - 1.5 = 4.5$. But then based on his or her reaction function, my adversary would want to produce $Q_2 = 6 - 0.5*4.5 = 6 - 2.25 = 3.75$. But if my adversary produced 3.75, then I would want to produce $Q_1 = 6 - 0.5*3.75 = 6 - 1.875 = 4.125$. But then if I produced 4.125, my adversary would want to produce $Q_2 = 6 - 0.5*4.125 = 6 - 2.0625 = 3.9375$. Can you see where this is going? It's taking us to both producing 4—the result we obtained above when we substituted one reaction function into another.

When both produce 4, we calculate the market price of 6 and the profits of each firm at 6 as above.

Suppose that the firms compete on price? They would compete down to marginal cost ($=2$). The market price would be 2, and the quantity demanded would be $P = 2 = 14 - Q$ or $Q = 12$. If each firm produced 6, its total revenue would be $2*6 = 12$, its variable cost would be $2*6 = 12$, its fixed cost is 10, and each would lose 10. Clearly, Cournot is a lot better (both make 6) than price competition.

What if the firms colluded and formed a cartel? The first seven columns of our spreadsheet showed us that if one firm served the whole market, that firm would price at 8, produce 6 units, and make a profit of 26. If the firms cartelized the market, they would split the output so that each made 3. Each would have total revenue of $8*3 = 24$, each would have total cost of $10 + 2*3 = 16$, and each would make a profit of 8. That is unambiguously better than the Cournot profits of each making 6. Why isn't the monopoly solution stable?

Think of firm 1's reaction function ($Q_1 = 6 - 0.5Q_2$). If I knew my adversary was going to produce 3, would I produce 3? My reaction function tells me the best thing for me to do is $Q_1 = 6 - 0.5*3 = 6 - 1.5 = 4.5$. If this happens,

$Q = Q_1 + Q_2 = 4.5 + 3 = 7.5$ Substituting this into the demand function gives $P = 14 - 7.5 = 6.5$. My total revenue is $6.5 \times 4.5 = 29.25$, my total cost $= 10 + 2 \times 4.5 = 10 + 9 = 19$, and my profit is 10.25, which is better than the 8 I get in our cartel. But if I have figured this out, what has my smart adversary figured out? The same thing. If my adversary produces 4.5, then together we've produced 9. Substituting 9 into the market demand function gives $P = 14 - 9 = 5$. Both of us have total revenue of $5 \times 4.5 = 22.5$, total cost of $10 + 2 \times 4.5 = 10 + 9 = 19$, and profit of 3.5. That's a lot less than the 6 each under Cournot. Unless we can legally sign binding contracts that have both of us making 3 units, the monopoly solution is not stable. Of course, neither is each producing 4.5 units. As shown above, if one firm produces 4.5, the other wants to produce 3.75, and so on, leading us to the Cournot solution of both producing 4.

EXCEL EXERCISE: STACKELBERG

In the Cournot exercise, the market demand was $P = 14 - Q_1 - Q_2$, each firm had a marginal cost (MC) of 2, average variable cost (AVC) of 2, and a fixed cost of 10. We developed each firm's reaction function, that is, $Q_1 = 6 - 0.5Q_2$ and $Q_2 = 6 - 0.5Q_1$.

Suppose that firm 1 has a first-mover advantage. That means that firm 2 will know how much firm 1 has made when firm 2 gets to move. Hence, to maximize their profit, firm 2 will make $Q_2 = 6 - 0.5Q_1$.

Having put itself in firm 2's shoes, firm 1 knows that firm 2 will produce $Q_2 = 6 - 0.5Q_1$. Firm 1 then substitutes for Q_2 in the market demand curve and determines its residual demand curve (i.e., what's left for firm 1 after firm 2 produces $Q_2 = 6 - 0.5Q_1$) is

$$\begin{aligned} P &= 14 - Q_1 - Q_2 = 14 - Q_1 - (6 - 0.5Q_1) = 14 - 6 - Q_1 + 0.5Q_1 \\ &= 8 - 0.5Q_1 \end{aligned}$$

Firm 1's total revenue is

$$TR_1 = P \times Q_1 = (8 - 0.5Q_1) \times Q_1 = 8Q_1 - 0.5Q_1^2$$

Firm 1's marginal revenue is

$$MR_1 = dTR_1/dQ_1 = 8 - Q_1$$

To maximize profits, firm 1 will set marginal revenue equal to marginal cost, that is,

$$MR_1 = 8 - Q_1 = 2 = MC_1$$

or

$$Q_1 = 6$$

Substituting $Q_1 = 6$ into firm 2's reaction function gives

$$Q_2 = 6 - 0.5 \times 6 = 6 - 3 = 3$$

Then $Q = Q_1 + Q_2 = 6 + 3 = 9$.

Substituting $Q = 9$ into the market demand curve gives

$$P = 14 - 9 = 5$$

Firm 1's total revenue = $TR_1 = P \cdot Q_1 = 5 \cdot 6 = 30$.

Firm 1's variable cost = $VC_1 = AVC_1 \cdot Q_1 = 2 \cdot 6 = 12$.

Firm 1's fixed cost = $FC_1 = 10$.

Firm 1's Profit = $\pi_1 = TR_1 - VC_1 - FC_1 = 30 - 12 - 10 = 8$.

Firm 2's total revenue = $TR_2 = P \cdot Q_2 = 5 \cdot 3 = 15$.

Firm 2's variable cost = $VC_2 = AVC_2 \cdot Q_2 = 2 \cdot 3 = 6$.

Firm 2's fixed cost = $FC_2 = 10$.

Firm 2's Profit = $\pi_2 = TR_2 - VC_2 - FC_2 = 15 - 6 - 10 = -1$.

Firm 2's Variable Profit = $V\pi_2 = TR_2 - VC_2 = 15 - 6 = 9$.

As you can see, it's good to go first since the profit 8 of firm 1 exceeds the simultaneous move profit of Cournot (6). Note also that it's bad to go second, since the profit of -1 of firm 2 is less than the Cournot profit of 6. If the fixed cost was sunk, firm 2 would stay in the market because the controllable cash flow is positive (15 in and 6 out). But if the firm had not yet committed the 10 in fixed cost, it would be a controllable cash flow, and the firm would not enter because cash flow out (16) would exceed cash flow in (15). That would then allow firm 1 to be a single monopolist and produce 6, price at 8, and make the profit of 26 as shown in the first seven columns of the spreadsheet exercise on Cournot.

How can we use a spreadsheet to get the above results? We have derived firm 2's reaction function in the Cournot spreadsheet exercise ($Q_2 = 6 - 0.5Q_1$). But suppose that we didn't know that but only the answers to all the what-if questions we asked in the Cournot exercise. Recall that those questions and answers were as follows:

If I produce Q_1	Then the optimal amount for my adversary to produce (Q_2) is:
12	0
10	1
8	2
6	3
4	4
2	5
0	6

Call up your spreadsheet. Enter 12 in cell A1, 10 in cell A2, and so on, until you enter 0 in cell A7.

Enter 0 in cell B1, 1 in cell B2, and so on, until you enter 6 in cell B7. Columns A and B are firm 2's reaction function.

Enter $=14-A1-B1$ in cell C1. Then click on the bottom right corner of cell C1 and drag your mouse down to cell C7. This will give you a column of the market price that goes with the corresponding sum of outputs from Columns A and B.

Enter $=A1*C1$ in cell D1. Then click and drag until cell D7. Column D will give the first mover's (firm 1's) total revenue.

Enter $=10+2*A1$ in cell E1. Then click and drag until cell E7. Column E will give the first mover's total cost.

Enter $=D1-E1$ in cell F1. Then click and drag until cell F7. Column F will give the first mover's profit.

Then search column F for the highest number (or let the spreadsheet find it for you by entering $=MAX(F1:F7)$ in cell F8). It will be 8 in cell F4. Reading leftward in row four, you will see in cell A4 that the first mover will produce 6 and in cell B4 that the follower firm will produce 3.

Enter $=B1*C1$ in cell G1. Then click and drag until cell G7. Column G gives the follower's total revenue.

Enter $=10+2*B1$ in cell H1. Then click and drag until cell H7. Column H gives the follower's total cost.

Enter $=G1-H1$ in cell I1. Then click and drag until cell I7. Column I gives the follower's profit. The follower does not get to choose the highest number in column I. Their destiny is sealed when the first mover chooses to produce 6. The follower's profit is found in cell I4 and equals -1 .

Enter $=10+I1$ in cell J1 and click and drag until cell J7. Column J gives the follower's variable profit. The follower gets variable profit of 9 in cell J4 when the first mover chooses output 6.

So we get the same results as the calculus using the spreadsheet.

CHAPTER 12

LEARNING OBJECTIVES

Making Strategy and Game Theory

Strategy Basics

Visual Representation

Solution Concepts

Equilibria

Dominant Strategies

The Nash Equilibrium

Strategic Foresight: The Use of Backward Induction

Repeated Games

Incomplete Information Games

Reputation Building

Coordination Games

Strictly Competitive Games

Summary

Problems

Excel Exercise: Game Theory

GAME THEORY

MAKING STRATEGY AND GAME THEORY

As we have seen, the managerial world is one of interactions with others. In that way, it mirrors life. We now offer a tool to help managers cope with the complexities of managerial life. In fact, it is a tool to help people cope with much of life. Let us explain. We can classify all managerial decisions as either strategic or not. Nonstrategic ones do not involve other people, so their actions need not be considered. For example, managers of a shipping company can generally map the most efficient shipping route without considering what other shipping companies are doing.

Strategic decisions are fundamentally different. Characterized by interactive payoffs, they require a different cognitive frame. Payoffs are **interactive** when the consequence of a manager's decision depends on both the manager's own action and the actions of others. For example, managers decide to enter a new market. Their payoff depends on whether others follow into the market.

So when managers ponder strategic decisions, they must explicitly consider what actions others will take. And optimal choices may change depending on managerial beliefs about others. But if only life were that simple: Others are doing the same thinking about you. And this is just the first link of the expectation chain. Did you ever change a decision because you thought others expected you to take it? Making a strategic decision is like looking at yourself in a hall of mirrors; except in strategy, there are others with you.

Interactive When the consequence of a manager's decision depends on both the manager's own action and the actions of others.

Game theory helps managers cope with their lives, not solve all their problems. The theory is that of mathematicians who tried to bring order to the complexity of life. Game theory, their organizing framework, can help managers better understand others strategically. Through its use, managers increase their ability to anticipate the actions of others. This ability, in turn, increases the payoffs of managerial decisions—almost as if managers can look into the future.

One rule that is basic to formulating strategy is the direct result of interactive payoffs: the lack of an unconditional optimal strategy. There is no best strategy for all situations; optimality is conditional on situational parameters, many of which are controlled by managers. Though a strategic situation challenges the decision-making skills of managers, it also offers opportunities for managers to change the parameters they control to increase firm payoffs. It is the duty of managers to recognize such opportunities and act accordingly. For example, changing a relationship to the long term from the short can alter the behavior of others (as we show later in this chapter).

Like any long-standing theory, game theory furthers managers' understanding of strategy on several levels. Gravity is a theory that most of us understand, although few can cite its mathematical formula. Game theory is similar in that it illustrates principles that, if followed, will lead to better decisions.

These principles are grounded in Taoist philosophy and were recorded over 2,000 years ago in a series of classic writings, such as *The Art of War*. They are the path to strategy because they should always be followed. The mathematics of game theory clarifies why managers need to follow these principles. Those who follow will make better decisions. The most important principle managers must remember is that they control their strategic environment. Because of interactive payoffs, actions by managers will induce others to change their behavior. As we said, optimality is conditional.

So even managers who simply follow the principles and do not solve for an equilibrium will increase their decision-making ability (though not as much as those who think more rigorously). This greater clarity comes from the visual identification and organization of game parameters.

Because we want to help managers cope better with managerial life, many of the situations we discuss combine conflict with mutual dependence. Such situations are common in the business world—price wars, negotiations, intrafirm relations—and skilled managers understand the relevant considerations.

STRATEGY BASICS

Before we play a game, we need to understand the rules; consider the game of poker. There are many ways to play poker; the rules determine the specific game being played. Game parameters, such as how many cards will be dealt, the betting procedure, and which hands are better than others, define the particular game

being played. The same is true for any strategic situation; the rules (parameters) define the game. Hence before managers act, they need to assess and understand the rules.

Sun Tzu said, “Strategy is the Tao of survival or extinction. One cannot but examine it.” Assessment is a basic principle. Although this should seem obvious to most managers, empirical evidence shows it is not practiced. For example, the most significant difference between good and bad negotiators is how they prepare for the negotiation. Those with better outcomes prepare more thoroughly.

Game theory provides an assessment framework to help managers. All game theoretic models are defined by a common set of five parameters. Other factors may influence managerial decisions, and game theory usually recognizes these. Here are the defining, common five factors:

1. **The players:** A player is defined as the entity making a decision; entities are either individuals or groups. The decisions of all players determine the outcome. These other entities are in the hall of mirrors with you, looking at the same situation as you but from their viewpoints. Models describe both the identities of the players and their number; changes in either can alter play. Identities are important because of the diversity of the universe. You need to know exactly who is in that hall of mirrors with you. (You cannot know who is playing without a scorecard.) For example, does your behavior change when you are with your parents rather than roommates? How about a total stranger? Most of us act differently if the individual (or group of individuals) with whom we interact changes. The model explicitly recognizes this, so it requires identification. Changes in the number of players can also alter strategies.

2. **The feasible strategy set:** Managers cannot anticipate or assess an action they believe is impossible. So only actions given a nonzero probability of occurring are assessed within a model. These actions constitute the feasible strategy set. Think of it as the potential behavior of others. Behavior not in this set is outside the analytic limits of a particular game model.

It is important for managers to think carefully about the strategy set. Any strategy they don’t consider, but is then played by others, puts managers at a strategic disadvantage. From a strategic view, being surprised is bad; it means that others are behaving in ways you did not anticipate. How do you know what to do if you have not assessed the situation?

3. **The outcomes or consequences:** Game models visually represent the intersection of the first two parameters as outcomes. Each player has a feasible strategy set (that is, behavior) comprising individual potential strategies. Each strategy of a player intersects all combinations of the strategies of others to form the outcome matrix. A particular outcome is defined by the strategy choice of each player. Think of the outcome matrix as a crystal ball containing all possible future states. After all players choose and play their strategies, the matrix identifies the

designated state as the future. Like a fortune-teller, game theory tries to predict which state will occur before it is actually realized.

4. The payoffs: A model assigns a payoff for each player to all outcomes. So the payoff for an outcome is expressed as a vector of individual payoffs, and each possible outcome has a corresponding payoff. A player's payoff is based on his or her preferences. An inherent assumption of game theory is that players are rational: They do not wish to harm themselves. Given the choice and all other things equal, they prefer a higher payoff to a lower one. This is another reason why player identities are important. Preferences are subjective: A payoff not highly valued by one player can be highly valued by others.

5. The order of play: Timing is important in both love and war. A model specifies the order in which players reveal their chosen strategies. Models are simultaneous if all managers reveal their strategies without knowing the strategies of others. Simultaneous play is not entirely time dependent. All players need not announce their decisions at precisely the same moment. It is more a matter of information. If all players commit to a strategy before learning the strategies of others, then the game is simultaneous. Nonsimultaneous games are by definition sequential. In any sequential game, the model specifies the order of play.

One way to measure the usefulness of a model is to examine how congruent it is with the real world. Here we summarize the mathematical framework to help managers decide whether it captures the nature of their world. Do managers consider the following in choosing a strategic action?

- How the outcome depends on their actions and the actions of others.
- The identities of others involved in the decisions.
- The order of play.
- How others will react to their decisions.
- The goal to achieve outcomes favorable to their preferences.

We believe most managers would answer these questions in the affirmative. That makes game theory relevant for the managerial world. Game theory formally analyzes what most managers intuitively consider when making strategy. Building a game model requires information managers already know. It asks for a finer partition of the information—a sharpening of managerial strategic focus. The theory's visual representations are tools to help managers with that focus.

VISUAL REPRESENTATION

Game models visualize interactive payoffs (outcomes) as the intersection of individual players' strategies. The representation of these payoffs takes one of two forms: matrix or extensive. The two represent the same information, although

Matrix form Form that summarizes all possible outcomes.

Extensive form Form that provides a road map of player decisions.

Game trees Games trees are another name for extensive form games and are akin to decision trees.

sequential games are more easily shown in the extensive form. The **matrix form** summarizes all possible outcomes; the **extensive form** provides a road map of player decisions.

Figure 12.1 represents a matrix form game of the following situation. Managers at two firms, Allied and Barkley, discover they both are planning to launch product development programs for competing products. They can choose to either keep spending at the currently planned level or increase it in hopes of speeding up product development and getting to the market first. Expected profits are a function of the expected development costs and revenues.

Let us see how the matrix addresses the five common parameters:

1. **Players:** There are two players, Allied and Barkley.
2. **Order of play:** Simultaneous—each must reach a decision without knowing that of the other.
3. **Feasible strategy set:** Each player can either maintain the current spending or increase it.
4. **Outcomes:** Because there are two players, and each has two strategic options, there are four possible outcomes.
5. **Consequences:** The payoffs are listed for each player within every possible outcome. The convention in game theory is to list the row player's (Barkley) payoff first in each cell and the column player's (Allied) second. So if Allied increases spending but Barkley does not, Allied's expected profit is \$3 million and Barkley's is \$2 million.

Extensive form games are also called **game trees**. These are akin to the decision trees we visit in Chapter 14 and are figuratively the same. The fundamental

FIGURE 12.1

A Two-Person Simultaneous Game

		Allied's strategy	
		Spend at current level	Increase spending
Barkley's strategy	Spend at current level	3, 4	2, 3
	Increase spending	4, 3	3, 2

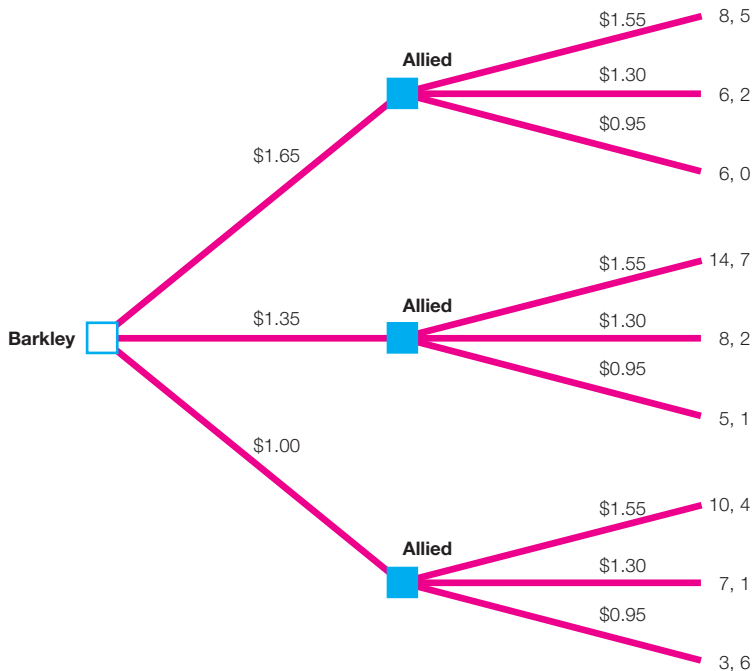
difference between a game tree and a decision tree is one of strategy. A game tree is strategic; a decision tree is not. Decision trees have no interactive payoffs; payoffs are a function of the single individual and nature.

Think of any extensive form game as a decision road map. Just start at the beginning and you cannot get lost. The extensive form gives more details than the matrix form by explicitly stating the timing of choices among players. Extensive form games represent the revelation of a player strategy with decision nodes. The node specifies the player's identity and the feasible strategy set (that is, behavior). The first node (decision) of the game is represented by an open square, all other nodes are shown by a solid square. Lines from each node represent the elements of the feasible strategy set. If other players reveal strategies later in the game, the lines lead from one node to another to show the order of play. If the player is the last to reveal his or her strategy, the lines lead from the node to a payoff schedule.

Figure 12.2 shows an extensive form game representing the following situation. Managers at Allied and Barkley must choose a pricing policy for the new product. They know the other will introduce a similar competing product. Because Barkley is expected to enter the market slightly sooner than Allied, Barkley

FIGURE 12.2

Allied-Barkley Pricing: Sequential



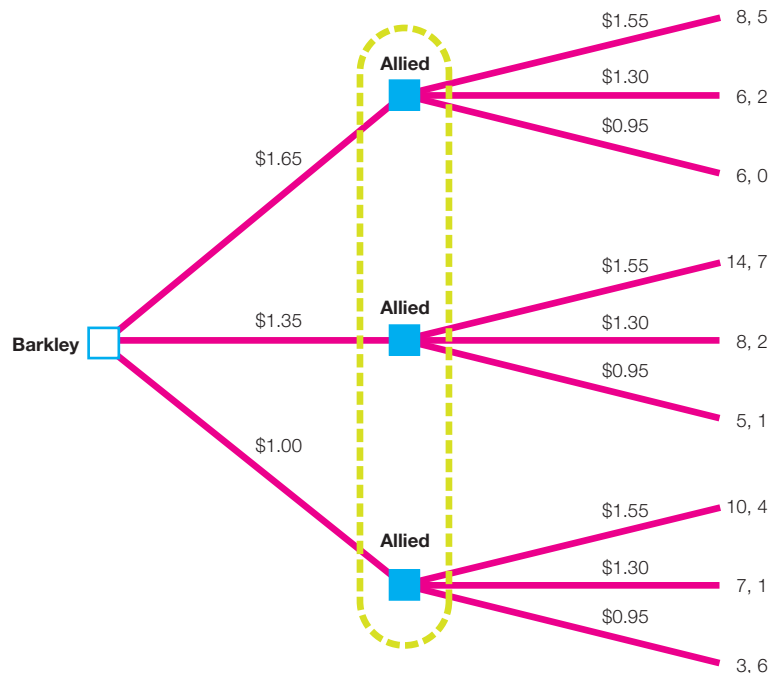
managers announce their price first (note the clear square). Managers will choose one of three prices: \$1.00, \$1.35, or \$1.65. Allied managers will reveal their price later. Because Allied is second to market, managers have possible price points of \$0.95, \$1.30, and \$1.55. Payoffs, which represent profits, are a function of costs and revenues.

The extensive form can also show simultaneous games. It does this with information sets. All simultaneous games are played with imperfect information; that is, when revealing his strategy, a player does not know the strategies of all others. This is the nature of simultaneous situations. Figure 12.3 shows the simultaneous version of the situation depicted in Figure 12.2. The only difference between the two figures is the dotted line drawn around the Allied decision nodes. It represents Allied's information set, or knowledge at the time it reveals its strategy. The dotted line signifies that Allied managers know they are at one of the three nodes—but not which one because Barkley managers have not revealed their strategy.

These two types of models show how game theory represents strategic situations. As you see, the information required is not voluminous. But it does require

FIGURE 12.3

Allied–Barkley Pricing: Simultaneous



some thought and reflection. The models provide a good organizing framework. They allow managers to communicate more clearly in a common language. The models are good for examining what-if scenarios—a favorite (and helpful) pastime of managers. If this was all they offered, we think the marginal benefits of learning them would far outweigh the marginal costs. But the best is yet to come. These models let managers anticipate the future by predicting the actions of others (and those of the managers themselves). It is important to understand the models' underlying strategic principles. This knowledge gives managers greater insight into human and corporate behavior.

SOLUTION CONCEPTS

How does game theory let managers see the future? It anticipates (correctly) the behavior of others. Using principles of the Taoist tradition, the theory of games adds formal ideas and parameters. How do models anticipate behavior? Unlike crystal balls and tarot cards, the theory provides details of the underlying logic of its predictive power.

Much of the theory's predictive power comes from solution concepts, which are basically rules of behavior. The theory anticipates behavior because it believes individuals act according to prescribed rules. Before we examine some of these rules, we need to understand the concept of an equilibrium.

EQUILIBRIA

Equilibria predict behavior in the following way. In an **equilibrium**, no player has an incentive to unilaterally change his or her strategy. This rule of behavior says little about how individuals arrive at this point; but once there, no one moves unilaterally. So game models predict individual behavior for each player and identify the future (outcome). The concept of an equilibrium holds this behavior in place.

Players do not unilaterally change behavior because they cannot increase their payoff. In an equilibrium, of all possible choices (the feasible strategy set), the present choice rewards players with the highest payoff if no other player changes behavior. In other words, conditional on the choices of others, we are doing the best we can (remember—payoffs are interactive). Thus we assume that behavior is directed by our preferences. We are rational in the sense that we do not want to hurt ourselves and accept a lower payoff. Hence an equilibrium is rational, optimal, and stable. A player's behavior is directed by a preference function. Each player tries to obtain the highest payoff possible, given the actions of others. Once an equilibrium is reached, no player has an incentive to unilaterally change behavior.

Equilibrium When no player has an incentive to unilaterally change his or her strategy.

STRATEGY SESSION: Managerial Use of Game Theory

The growing use of game theory by managers is reflected by articles in the business press that analyze business decisions within a game theory context.

Paul Kerin of *The Australian* uses game theory to analyze a decision by BHP Billiton CEO Marius Kloppers to announce support of a carbon tax in Australia. Kloppers is known to use game theory to analyze acquisitions and other business decisions. Kloppers' support for the tax caught many by surprise. After all, a carbon tax is generally not supported by the CEO of a company that is a major supplier of carbon-related raw materials.

According to Kerin, the key to understanding why Kloppers might support a carbon tax is to think about what alternatives the government might consider. One prominent alternative is the emissions trading scheme (ETS). An ETS creates a market where managers can purchase and sell permits for emissions of specified pollutants. Managers of firms which produce lower levels of emissions than specified can sell some of their extra "credits" to firms that produce emissions above the regulated levels.

If everything was equal, a tax should produce the same results as an ETS. But, there are actually fundamental differences between the two in how they operate. A critical difference is the price setting mechanism. In the ETS market, the price is set by the

market. For the carbon tax, the price is set by the government. Clearly, it is easier for Kloppers to influence the political process than try to interfere with the market process. In Kloppers' favor, his lobbyists would be dealing with a new, minority-led government.

From Kloppers' view, another disadvantage of the ETS market is that it reveals information about the marginal cost of cutting emissions. In fact, if efficient, the market-clearing price clearly reflects the marginal cost of abatement. Firms with high marginal costs of abatement will purchase permits or credits while those with low marginal costs will be net sellers. If the government institutes the carbon tax, the government receives no information about the marginal costs of emissions reduction. In fact, the only way they could determine this is to ask managers at each polluting company what their marginal cost is and then identify the firm with the lowest marginal cost. Of course, managers have no incentive to reveal their true marginal cost.

Kloppers also considered that, in the recent Australian election, the Green Party won a significant number of seats. They favor an ETS. Kloppers hoped to forestall their efforts, hence his support for the carbon tax. Supporting the carbon tax also allows Kloppers to truthfully say he is in favor of environmentally-friendly regulations.

DOMINANT STRATEGIES

One way to tame the complexity of strategic thinking is to make it less strategic. What if managers act without regard to the actions of others? This takes managers out of the hall of mirrors and puts them in front of one mirror. Looking back at them are the only people responsible for optimizing payoffs in nonstrategy land: themselves. In these situations, if managers do not choose an optimal strategy, they have only themselves to blame.

But mathematicians cannot ignore the interactivity of life; and life suggests there are times when managers possess an option that is strategically strong rela-

At the time of his announcement some accused Klopers of not protecting the interests of shareholders. Clearly, those who accused Klopers did not understand the strategic nature of the situation. Forced to choose between two paths, Klopers chose the one where he could maintain more control. Frankly, he chose the path that increased firm value.^a

On October 13, 2009, Dennis Berman of *The Wall Street Journal*, analyzed the proposed Comcast-NBC deal using a game theory model. Developed by Dr. Bueno de Mesquita, the model analyzes a merger situation and predicts outcomes. In analyzing the situation, the model recognizes the complexities of managerial behavior. For example, CEOs may pay a high price to acquire because it is in their interests, not those of the shareholders. The model also accounts for perceived negotiating power among the parties involved.

As a metric of the current state of managerial behavior, de Mesquita has programmed the model to assume all parties are purely self interested. This must be accurate because, according to declassified CIA documents, when CIA personnel “back-tested” predictions, they were correct 9 of 10 times.

One issue the model looked at was the anticipated behavior of a third firm, Vivendi. A large obsta-

cle to the Comcast-GE deal was that Vivendi held a 20% stake in the GE Universal venture. They could sell the shares back to GE or use them to create an IPO. This threat certainly affected the share price negotiation between Vivendi and GE.

Model inputs consisted of both qualitative and quantitative data. For example, the model asked for input on managerial personalities, motives, and bargaining power. Valuation scenarios were created and analyzed.

After conducting the analysis, the model’s most likely scenario was for Vivendi to sell its stake to GE at a price slightly below the expected valuation. On December 3, 2009, GE and Vivendi reached a purchase agreement. GE would purchase the 20% for \$5.8 billion, slightly lower than the \$6 billion many observers expected.^b This shows the power of game theory to correctly anticipate the future decisions of managers. Used well, game theoretic reasoning is a valuable managerial tool.

^a Paul Kerin. “A Game Theorist Plays a Trump Card,” *The Australian*, September 20, 2010, p. 26.

^b Dennis Berman. “In NBC Deal, Learn from Game Theory,” *The Wall Street Journal*, October 13, 2009, p. C1.

tive to all others. Mathematicians represent these beliefs as **dominant strategies**. A dominant strategy is one whose payoff in any outcome is higher relative to all other feasible strategies.¹ Managers choose dominant strategies to optimize their expected return. Although the strategy choices of others still affect managerial payoffs, thinking about others will not change the managerial decision. A dominant strategy returns a higher payoff than any other strategy across all possible outcomes. Managers should always choose a dominant strategy if it is available.

It is easier to visually represent dominant strategies using the matrix form. Look at Figure 12.1 again. Managers at Allied have a dominant strategy to maintain the current spending level (left column). If managers at Barkley also maintain

Dominant strategies A strategy whose payout in any outcome is higher relative to all other feasible strategies.

1. A strategy is weakly dominant if it does at least as well as any other strategy for some outcomes (it is tied with another for the highest payoff) and better than any strategy for the remaining outcomes. Even if a strategy is only weakly dominant, a player would choose it.

current spending, Allied managers receive a payoff of \$4 million (in blue); if Barkley managers increase spending, Allied managers earn \$3 million (in blue). Barkley managers also face a dominant strategy: They will increase spending. If Allied managers spend at the current level, Barkley managers earn \$4 million; if Allied managers increase spending, Barkley managers earn \$3 million.

We can now predict that Allied managers will maintain their current spending and those at Barkley will increase it. We have a dominant strategy equilibrium. Why will this outcome prevail? Because each manager has a dominant strategy and each is always better off playing it. Why won't they change their strategy? Any change in strategy will lead to guaranteed lower payoffs. Domination is the minimum hurdle required of rationality. If you choose dominated strategies when you know dominant ones exist, you are hurting yourself.

Dominant strategies are the stress reducers of the strategic world. They ease the mental cost of decision making and simplify the analytic process. With them, managers can ignore the actions of others. Given the hectic schedules of managers, they are a great time-saver. For example, in Chapter 13 we discuss auctions. Auction design can determine whether a dominant strategy exists for a particular auction. In auctions with no dominant strategy, managers must consider the bids of others in choosing their bids. In ones with a dominant strategy, managers need not. Good strategic managers understand the difference and do not waste time thinking about something irrelevant to a decision.

Not surprisingly, it is dominant to first look for dominant strategies in matrix games. Even if a game is not solvable through dominance, this process eliminates outcomes mapped to dominated strategies; we essentially reduce the set of playable strategies. Recall that no player should ever play a dominated strategy because he or she is always better off playing the dominant one. So we can eliminate dominated strategies from consideration. We never reach those outcomes because no dominated strategy is played. Visually, our outcome matrix is reduced by rows or columns. Mentally, the analysis is simplified because we consider fewer outcomes.

More important strategically, when managers eliminate a strategy, they reduce the set of possible outcomes. This in turn may change a formerly nondominated strategy into a dominated one. This iterative process can proceed until each player is left with only one playable strategy (that is, the situation is dominance solvable).

Figure 12.4 is the matrix form of the game shown in Figure 12.2. Managers can always model strategy as either an extensive or a matrix form game. Each game form is linked to a particular game of the other form with identical players, outcomes, and payoffs. They are like fraternal twins who act identically.

We now look at the strategies for Allied and Barkley managers. Barkley has a dominated strategy: \$1.35 dominates \$1.00. For any Allied strategy, Barkley managers earn more by pricing at \$1.35. If Allied managers price at \$0.95, \$1.35 earns 5 relative to 3 for \$1.00. If they price at \$1.30, the ratio is 8 to 7; and if Allied managers to price at \$1.55, the ratio is 14 to 10. Barkley managers should never choose

FIGURE 12.4

Matrix Form Representation of Figure 12.2

		Allied's pricing strategies		
		\$0.95	\$1.30	\$1.55
Barkley's pricing strategies	\$1.00	3, 6	7, 1	10, 4
	\$1.35	5, 1	8, 2	14, 7
	\$1.65	6, 0	6, 2	8, 5

\$1.00, so we need not further consider its outcomes. This effectively reduces the figure's matrix to that shown in part A of Figure 12.5.

Because we eliminated Barkley's \$1.00 strategy as a playable strategy, Allied managers now face a dominated strategy—two in fact: \$1.55 dominates both \$0.95 and \$1.30. Hence the matrix is reduced to that shown in part B of Figure 12.5. Allied managers have only one playable strategy: charging \$1.55. Therefore, Barkley managers now face another dominated strategy. Because 14 is greater than 8, Barkley managers will charge \$1.35. So this game has a dominant strategy equilibrium: Allied managers will price at \$1.55 and Barkley managers will price at \$1.35.

Although both this game and the one represented in Figure 12.1 have dominant strategy equilibria, they are solved using different degrees of rationality. Most people view rationality as dichotomous: If you are not rational, you are irrational. But rationality has continuous measures. One such measure is the degree of rationality required to reach an equilibrium. Because of interactive payoffs, we have to worry about not only our own rationality but also the rationality of others. So the degree of rationality measures the number of conjectures required to reach an equilibrium.

For example, in Figure 12.1 managers can solve the game simply by knowing they are rational. If you are rational, choose your dominant strategy. No matter what others do, that is the best response. But this single degree of rationality is not sufficient to ensure that managers reach the equilibrium in Figure 12.4. When dominance is of the iterative variety, managers need to consider the rationality of others.

Let us view the game from the decision viewpoint of Barkley managers. "Well, we know we are rational. We see that \$1.00 is dominated, so we will not price at \$1.00. Fine; what are our beliefs about the behavior of Allied managers? Will they

FIGURE 12.5

Iterative Dominance

A. Barkley's \$1.00 strategy is eliminated.

		Allied's pricing strategies		
		\$0.95	\$1.30	\$1.55
Barkley's pricing strategies	\$1.35	5, 1	8, 2	14, 7
	\$1.65	6, 0	6, 2	8, 5

B. Allied's \$0.95 and \$1.30 strategies are eliminated.

		Allied's pricing strategies	
		\$1.55	
Barkley's pricing strategies	\$1.35	14, 7	
	\$1.65	8, 5	

believe our price of \$1.00 is dominated and we will not price there? If so, do Allied managers realize that pricing at \$1.55 is a dominant strategy?"

In matrix form games, we measure the degree of rationality by the rounds of iterative dominance needed to reach the equilibrium. The game in Figure 12.4 requires three rounds. It is our experience that many people act rationally for games with a few degrees of rationality, but few people understand games of many degrees.

Life is complex. Domination principles help make managerial life easier. But few games have a dominant strategy equilibrium; interaction with others is generally more complicated. How can managers anticipate behavior in games without a dominant strategy equilibrium?

THE NASH EQUILIBRIUM

That is the question mathematician John Nash asked himself in the early 1950s. His answer is our most widely used solution concept: the Nash equilibrium. Similar to dominance, Nash developed guidelines for behavior that are rational,

optimal, and stable. He specified behavior if players lacked a strategy that dominated all others. Here is the intuition underlying Nash's ideas.

A player's objective remains the same whether he has one or more playable strategies: to maximize the payoff. If he must choose among playable strategies, he selects the optimal one. And because payoffs are interactive, he must choose conditional on what he thinks of others. So what rule are the others following? If they are the same (rational), their goals are identical: to maximize their payoffs relative to what others will choose. Hence each will choose the strategy that maximizes his or her payoff, conditional on others doing the same. That is John Nash's prescribed behavior for players with more than one playable strategy.

Nash's concept is more transparent in its mathematical form. Let each of N players identify a strategy s_i^* , where $i = 1, 2, 3, \dots, N$. An outcome represents an array of strategies $s^* = (s_1^*, s_2^*, s_3^*, \dots, s_N^*)$. Let $B_i(s^*)$ be the payoff to player i when s^* is chosen, with i being any player, $i = 1, 2, 3, \dots, N$. Then a Nash equilibrium is an array of strategies such that

$$B_i(s_1^*, s_2^*, s_3^*, \dots, s_N^*) \geq B_i(s'_1, s_2^*, s_3^*, \dots, s_N^*) \quad \text{for all outcomes}$$

The left side of this equation states the existence of an outcome(s), defined by the array of player strategies, where all have a best response to the best responses of others. The right side states that if any player unilaterally changes strategy, she realizes a lower payoff; that is, she chooses a dominated outcome. For a complex and messy process, Nash devised an elegant solution. It treats all with equal rationality and so limits behavior. Dominance is still present, although now conditional. The solution exists for all games with a finite number of players and outcomes.

Dominance is unconditional in dominance-solvable games. We need not speculate about the behavior of others because it makes no difference. But anything less than unconditional dominance requires anticipation. And because payoffs are interactive, this anticipation requires a common vision with others. So Nash prescribes a behavioral rule. Maximize your payoff, conditional on all others doing the same. A Nash solution is dominant, conditional on this rule being followed.

Recall that an equilibrium needs to be rational, optimal, and stable. Nash's solution is rational in the sense that all players follow the prescribed behavior. It is optimal in that all try to maximize their payoffs. And it is stable because no player can unilaterally change strategy and realize a higher payoff.

Figure 12.6 illustrates the following. The numbers represent profits (in millions of dollars). Recall that Barkley entered the market first, followed by Allied. Both firms must now introduce new products. Each can choose one product of several; but because of financial constraints, only one can be supported. Managers at both firms understand this. Their choice to introduce a product is conditional on how they think the other will behave. Nash says that we all behave identically: We maximize payoffs conditional on others doing the same. We will change behavior to obtain a higher payoff but not a lower one.

FIGURE 12.6

New Product Introduction

		Allied		
		Product alpha	Product beta	Product zeta
Barkley	Product lambda	4, 6	9, 8	6, 10
	Product pi	6, 8	8, 9	7, 8
	Product sigma	9, 8	7, 7	5, 5

Look at Figure 12.6. Remember our decision rule: Look for dominated strategies. This is quickly done. Confirm that neither firm has a dominated strategy. Now use the following algorithm. For each strategy, indicate the behavior of the other. For example, if Barkley managers know (with certainty) that Allied will introduce alpha, what will they do? Barkley managers receive 4 if they introduce lambda, 6 if pi, and 9 if sigma (the first numbers in each cell of the product alpha column). Because 9 is the highest of the three payoffs, Barkley managers will introduce sigma if they know Allied will introduce alpha. Write a B in the alpha–sigma outcome. Do the same for strategies beta and zeta. Now follow the same procedure for Allied (the sequence of players makes no difference). For example, if Barkley introduces sigma, what should Allied managers choose? If Allied managers produce alpha, they receive 8, if beta 7, and if zeta 5. So Allied managers will introduce alpha if they know Barkley will introduce sigma. Mark an A in this outcome. Do the same for strategies lambda and pi. The resulting matrix is illustrated in Figure 12.7.

Any cell with an A and a B is a Nash equilibrium. In this game the Nash solution is for Barkley to introduce sigma (and receive 9) and Allied to introduce alpha (and receive 8). Let us understand why this outcome is predicted by Nash. An A or B represents a conditional dominant strategy—a best response to a specific strategy of others. A Nash equilibrium is a meeting of the best responses—an outcome where all play conditionally dominant strategies. Beyond the reach of an individual, it is attained by the group.

Each player acts in his or her own best interest and maximizes payoffs. A player not choosing the Nash strategy is playing a dominated one (assuming others play their Nash strategies). Hence no player has an incentive to unilaterally change behavior. For example, as long as Allied produces alpha, Barkley managers want to produce sigma and receive \$9 million. If they produce pi, they receive only \$6 million, and if

FIGURE 12.7

New Product Introduction with Other's Behavior

		Allied		
		Product alpha	Product beta	Product zeta
Barkley	Product lambda	4, 6	(B) 9, 8	6, 10 (A)
	Product pi	6, 8	8, 9 (A)	(B) 7, 8
	Product sigma	(B) 9, 8 (A)	7, 7	5, 5

lambda, only \$4 million. Allied managers face the same scenario: lower payoffs for any change in behavior. Interactive payoffs hold the two hostage to each other.

Recall our example of network economics in Chapter 6. Building a transportation network can be expensive but it does have strategic value. A new entrant has a very difficult task because shippers want to call one carrier who will take any package anywhere. As Fred Smith (the CEO of FedEx) once said: "You don't want to have a yellow phone to call Houston, a green phone to call London, a blue phone to call Los Angeles. . . . You want one color phone to call anywhere. You cannot enter this market piecemeal with just New York to Chicago service when an incumbent can go from New York to anywhere." This is a lesson that successful international giant DHL learned when they attempted to enter the U.S. package freight market. They entered in 2003 by buying Airborne, which had a less-than-complete network that they operated poorly. DHL could not provide the full service that the competitors could. DHL tried to counteract this disadvantage by offering cheaper rates, but this strategy was unsuccessful. The only other entrant into this market has been UPS, but they could piggyback on their already existing and very efficient ground network. Building the network first by FedEx initially lead to a monopoly in the overnight market, then a duopoly when UPS entered. The U.S. Postal Service provides a third player in this market but their network is mandated (and as we are painfully aware, the USPS is hemorrhaging money).

It should be noted that FedEx entered the ground transportation market when interstate motor carriage was deregulated in 1980 and further when intrastate motor carriage was deregulated in 1995. While UPS has integrated its ground and air network, FedEx runs two separate networks. But to this day, FedEx still has the largest market in the air part of the business (where it entered first) and UPS still has the largest market in the ground part of the business (where it entered first).

Let's look at how DHL competed with FedEx and UPS when DHL entered the U.S. domestic market for small package delivery (basically packages weighing up to 150 pounds—average weight much less) in 2002. DHL was (and still is) a powerhouse in international small shipments. FedEx and UPS were firmly established (and highly thought of) by U.S. domestic customers and both serve extensive international markets (just not as well as DHL).

Suppose there are two customers, A and B, of domestic small shipments. If they use FedEx and or UPS, they both profit by 100. Suppose if DHL could actually replicate their international service in the U.S. domestic market, each customer would benefit by 120 if they switched to DHL. However, if one customer used FedEx/UPS and the other used DHL, each customer would have a profit of 90 (the splitting of the market leads to less economies of scale, hence higher unit cost, and hence higher rates). The situation appears in the game theory matrix (Figure 12.8) below.

Using bolds to indicate Customer A's best choice *given* Customer B's choice and underlines to indicate Customer B's best choice *given* Customer A's choice shows two Nash equilibria: the northwest corner with both customers using their current carriers FedEx/UPS and the southeast corner with both customers using DHL. Clearly the southeast corner Nash is superior.

But before DHL entered the market, FedEx/UPS had locked the customers into their respective systems through long term contracts or through inertia (I've always used them). If both shippers would communicate with each other, they would mutually move to the southeast Nash (since $120 > 100$). But neither customer would *unilaterally* leave their current provider (because $100 > 90$). The difficulty is that there are basically thousands and thousands of system users. So although two can easily say to one another: I'll switch if you switch, for thousands to do so would be very difficult to coordinate. Of course, here's where national media and trade press can help spread the word about the benefits (sort of a Consumer Reports on the small shipments industry).

FIGURE 12.8

Payoffs for Each Customer From Using FedEx/UPS and/or DHL

		Customer B	
		Use FedEx/UPS	Use DHL
Customer A	Use FedEx/UPS	100, <u>100</u>	90, 90
	Use DHL	90, 90	120, <u>120</u>

Is there anything that DHL as a second mover could do by itself to break the stranglehold that FedEx/UPS's first mover advantage had over them? One way is to up the 90 payoff a customer gets from using DHL when the other customer uses FedEx/UPS so that it's over 100 (say 100^+). This could occur by offering price discounts for DHL services (and DHL did lower prices).

Suppose that a DHL lowering of prices changed the game matrix of Figure 12.8 to that of Figure 12.9 below.

Using the bolds and the underlines as in Figure 12.9 shows now that both Customer A and Customer B have a dominant strategy to switch to DHL. They don't need to coordinate; they just all do what's unilaterally best for themselves and they all switch. The first mover advantage of FedEx/UPS is broken.

If only that was the story for DHL. Alas, it is not. Their service level was so poor that their perceived payoffs to customers did not reach 100^+ even with their rate cuts. Thus the matrix result continued from Figure 12.8. DHL could never again much of a market share in the U.S. domestic small shipments market (perhaps 10% at best and that from customers who cared about price and not service) and lasted only five years in the U.S. market (each year proclaiming that next year they would be profitable and each year losing money which approximated \$10 billion from 2003–2008). It is interesting to note that the other major international player (TNT) has shown no interest in entering the U.S. domestic market (presumably because of the cost of building a network and overcoming the first mover advantages of FedEx and UPS).

Networks may be one way or two way (as the phone and FedEx networks above). The pipeline business (oil, oil products, natural gas, water, sewage) are one way networks. Oil flows from A to B in the pipe but there is no return trip. There are limited benefits to existing customers from adding more customers to the current system (unless the current system is undersubscribed). There are some advantages to customers and the pipeline when more customers moving the same product join.

FIGURE 12.9

Payoffs for Each Customer From Using FedEx/UPS and/or DHL after DHL Lowers Rates

		Customer B	
		Use FedEx/UPS	Use DHL
Customer A	Use FedEx/UPS	100, 100	100 ⁺ , <u>100⁺</u>
	Use DHL	100⁺ , 100 ⁺	130⁺ , <u>130⁺</u>

This is because the various customers product can be co-mingled and the pipeline doesn't have to add separators to keep different products segregated. There are disadvantages from having too many customers as congestion may develop in the network.

If the number of customers and their flows can be determined *before* the pipeline is constructed, then there are major advantages of adding new customers since a doubling of the pipeline diameter (a major input to the production of the pipeline's services is the amount of steel necessary to construct the pipeline) quadruples the area of the pipeline (the output that the pipeline can produce). This leads to major economies of scale, lowering the unit cost, and potentially lowering the rate. But with the system already in place, congestion raises its ugly head. Think of congestion on existing road networks and airports at rush hour, website crashes when too many people want to access a site at the same time, power failures or brownouts when the temperature skyrockets.

Not only do networks provide the network positive externalities shown above with respect to connectivity, but they can also provide network complementarities. Think of the development of Skype for use on the Internet. Just as with phone service, if there was no one to Skype with, the software would be not valuable. But as the Internet expands, more services will be developed. Here's just a few that have taken hold: Netflix DVD streaming, EBay auctions, online shopping, online dating services, etc, etc.

Think of other network complementarities that have developed because of networks. Electrical networks have generated a vast industry that produces all sorts of electrical appliances. The highway network (which was urged on the government by the auto industry) has grown the automobile industry and all of its complementary industries, e.g., tires, service stations, car detailers, roadside motels, AAA, a pine tree air freshener industry, etc. And as mentioned above, the Internet has lead to a software development industry. The list goes on.

Just one more example to bring the concept of network externalities home. Think of the value of the two way network if you run a dating service. It's hard to sign up the first customer because they'll ask: Who is there for me to date? But when lots are signed up, there are lots times lots minus one potential matches (and that's really a lot). The old method was tedious. You went to a bar and tried you luck or you placed a classified ad and waited for a response (and then you met up in a bar). But now you enter who you are and what you're searching for and you're matched with others who are "compatible." People, of course, lie on these questionnaires (especially about age, height, weight, good looks) but the number of "hookups" are non-trivial and some of the guessing is taken out of the search process. And, of course, the larger the number of subscribers to the dating service (members of the network), the higher is any given subscriber's probability of finding Ms. or Mr. Right. This same principle applies to things like online auctions. The more people that want to sell an item called X, the better it is for buyers of the item and the more people that want to buy a unique item Y, the better it is for the seller of that item.

STRATEGIC FORESIGHT: THE USE OF BACKWARD INDUCTION

A Taoist saying is, “Good strategists take care of the great while the great is small.” Good managers use strategic foresight. The ancient Taoists called this the ability to see what others could not see. We define **strategic foresight** as a manager’s ability to make decisions today that are rational given what is anticipated in the future. For example, a manager builds extra capacity today because she believes (correctly) that demand will increase in the near future. Strategic foresight is a principle that should always be used. Remember that the decisions of today can never affect your past—only your future. In decision making you always want to look forward. Using strategic foresight also helps managers understand that decisions have both short- and long-term consequences.

Game theory formally models strategic foresight through what is called **backward induction**. In game theory we use backward induction to solve games by looking to the future, determining what strategy players will choose (anticipation), and then choosing an action that is rational, based on these beliefs. Backward induction is most easily seen in extensive form games because of the ability to map the choices of players.

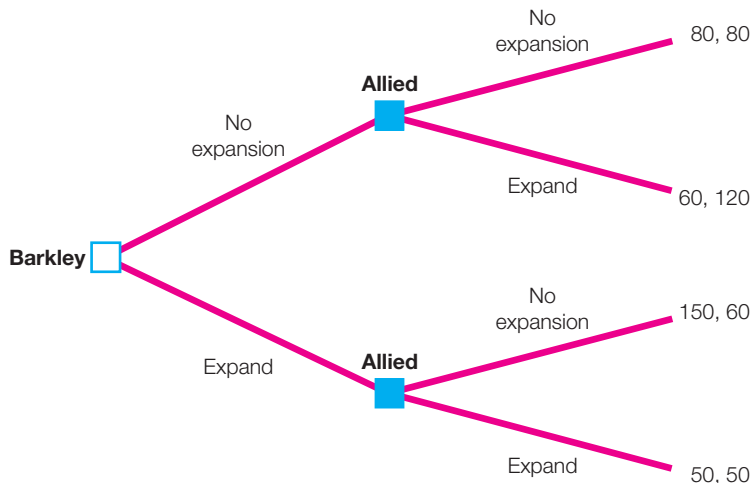
Figure 12.10 shows a game in extensive form. Recall from Figure 12.7 that Barkley managers have chosen to introduce sigma, and Allied managers have chosen to introduce alpha. They must now decide whether to expand their product lines. Either firm’s managers have the choice to expand or not. Barkley is the

Strategic foresight A manager’s ability to make decisions today that are rational given what is anticipated in the future.

Backward induction Used in game theory to solve games by looking to the future, determining what strategy players will choose (anticipation), and then choosing an action that is rational, based on these beliefs.

FIGURE 12.10

Allied–Barkley Expansion Decision



market leader, so assume its managers will reach their decision first. After seeing the decision of Barkley managers, those at Allied decide whether to expand. Payoffs for the four possible outcomes are given in Figure 12.10.

Let us see how a manager with strategic foresight can use backward induction to solve this game. The farthest left decision node represents the decision of Barkley's managers to expand or not. If they decide to expand, Allied's managers face the situation represented by Allied's bottom decision node. If Barkley does not expand, Allied faces the situation represented by the top decision node. Barkley managers use strategic foresight. They want to make a decision today that maximizes their payoff, given their vision of the future. Barkley managers realize that if they expand, Allied will receive a payoff of \$50 million if it expands and \$60 million if it does not.

Hence Barkley managers anticipate that if they expand, Allied will not. What if Barkley decides not to expand? Allied receives a payoff of \$120 million for line expansion and \$80 million if it does not expand. Thus, Barkley managers anticipate that Allied will expand if they choose not to. So if they expand, they anticipate a payoff of \$150 million (because Allied will not expand). If Barkley does not expand, its managers should anticipate a payoff of \$60 million (because Allied will expand). Since \$150 million is greater than \$60 million, the managers at Barkley know they should expand their product line.

To use backward induction we must come back from the future. We anticipate the future actions of others and then choose actions that are rational, conditioned on our expected behavior of others.

Backward Induction and the Centipede Game

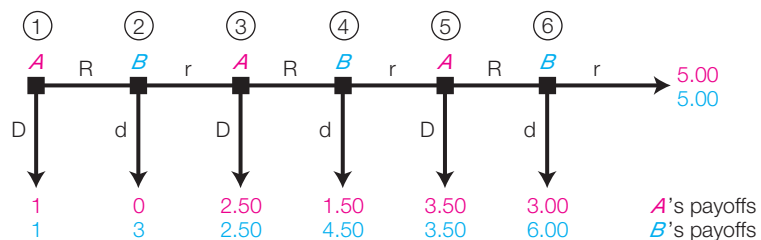
The usefulness of backward induction in strategic thinking is clearly shown in the simple **centipede game**. Many studies have used this game to study whether subjects use and understand backward induction. The game is shown in Figure 12.11.

Two players (*A* and *B*) participate in this sequential game. Player *A* moves first and can choose either down (*D*) or right (*R*). If player *A* chooses *D*, the

Centipede game A sequential game involving a series of decisions that shows the usefulness of backward induction in strategic thinking.

FIGURE 12.11

The Centipede Game



game is over and both players receive a payoff of \$1. If player *A* chooses *R*, then player *B* faces a similar choice. She can choose *d* or *r*. If player *B* chooses *d*, the game is over; *A* receives a payoff of \$0, and *B* receives a payoff of \$3. If player *B* chooses *r*, the game continues, and player *A* chooses either *D* or *R* again. The game continues, until one player chooses down or player *B* is asked to choose for a third time. At this point, if player *B* chooses *d*, *A* receives \$3 and *B* receives \$6. If player *B* chooses *r* at this point, both players receive \$5. Look at Figure 12.11 and assume you are Player *A*. What strategy would you choose?

We solve the game using backward induction. The game is actually a series of six decisions. Player *A* chooses at stages 1, 3, and 5; player *B* chooses at stages 2, 4, and 6. We need to go to the end of the game and work backward from the future. Look at stage 6: Player *B* can choose down and receive \$6 or choose right and receive \$5. Because \$6 is greater than \$5, we anticipate that player *B* will choose down. Move backward to stage 5 because we now know the future. Player *A* faces the following. If *A* chooses right, he knows (anticipates correctly) that *B* will choose down, giving *A* a \$3 payoff. Or *A* can choose down and receive a payoff of \$3.50. Because \$3.50 is greater than \$3, *A* will choose down at stage 5. What should player *B* choose at stage 4 knowing this? *B* can choose down and receive a payoff of \$4.50 or choose right and receive a payoff of \$3.50 (because *B* anticipates that *A* will choose down at stage 5). Because \$4.50 is greater than \$3.50, player *B* will choose down at stage 4. We move backward to stage 3. Player *A* can choose down and receive a payoff of \$2.50 or choose right and receive a payoff of \$1.50 (since *A* anticipates that *B* chooses down at stage 4). Because \$2.50 is greater than \$1.50, player *A* will choose down at stage 3. We are now at stage 2. Player *B* can choose down and receive a payoff of \$3 or choose right and receive a payoff of \$2.50 (since player *A* will choose down at stage 3). So player *B* will choose down at stage 2. Finally we find ourselves in the present; this is decision time. At Stage 1, player *A* can choose down and receive \$1 or choose right and receive \$0 (because *B* will choose down at stage 2). Player *A* will choose down at stage 1. This is the only rational choice, given our view of the future. Player *A* will choose down at stage 1, and both players will realize a payoff of \$1. What the ancients called foresight, game theorists model as backward induction.

Now comes the real question. How do subjects behave while playing the game? Initially few subjects appear to use foresight (or they have a distorted view of the future). Subjects appear to focus on the growing size of the payoffs and try to move down this path. At some late stage, either they choose down or the person they are playing with chooses it. The next time they play, most subjects tend to choose down at an earlier stage (especially those whose partners chose down in the earlier game). By the third or fourth play sequence, most *A* players are resigned to the fact that they should choose down at stage 1. They do so hesitantly because they still see the path of greater payoffs. But they also know the future. Experience has shown the wisdom of backward induction.

The Credibility of Commitments

Backward induction has many uses. One is to test the credibility of commitments. From threats to promises, epithet slinging to love's rhapsody, we want to know whether we should believe others. When facing these situations, always check credibility first. That is a dominant strategy. Consider only credible commitments. A commitment is **credible** if the costs of falsely sending one are greater than the associated benefits. Managers at a company who proclaim its product is best are not credible. There is little cost to proclaiming this, and associated benefits are high. The managers can make that claim credible by offering a product warranty. A warranty increases the commitment cost (if it is falsely sent). There are many uses for such costs.

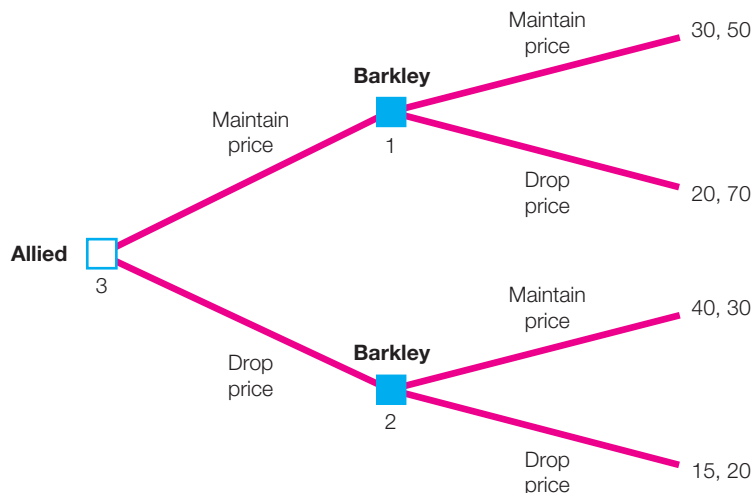
Credible commitment When the costs of falsely sending a commitment are greater than the associated benefits.

Consider the following. Recall that Barkley managers expanded their product line but Allied managers have not. Allied's managers decide to counter Barkley's product line extension by dropping the price of their product. However, they are concerned that if they drop their price, Barkley managers will follow with a price cut of their own. In fact, Barkley's managers told a common supplier of both firms that if Allied drops its price, they will drop theirs. What should Allied managers do?

Allied's managers first must consider whether Barkley's threat to drop its price is credible. They can do so by looking at Figure 12.12 and solving the game using backward induction. Barkley managers can either keep their price high or drop it. Allied managers have the same two strategies available to them. What happens if

FIGURE 12.12

Does Barkley Have a Credible Threat?



Allied managers drop their price? Those at Barkley can either keep their price high and earn \$30 million or drop their price and earn \$20 million. Because \$30 million is greater than \$20 million, Barkley managers should keep their price high if Allied managers reduce their price. What if Allied maintains its price? Barkley managers could keep their price high and earn \$50 million or drop their price and earn \$70 million. Hence Barkley managers will drop their price if Allied managers maintain their high price. Given that Allied managers can anticipate these actions from those at Barkley, how should Allied managers decide? They should drop their price. Clearly those at Barkley will not drop their price. The threat by Barkley managers to lower price is not credible, so it should be ignored. If forced to carry out the threat, Barkley managers will refuse. To do so would cost them \$10 million in lost profit.

The equilibrium just described is a subgame perfect equilibrium. A **subgame** is defined as being a segment of a larger game. The subgames are marked in Figure 12.12. As you can see, within the overall game, the three constitute three separate subgames. In repeated games, all subgame perfect equilibria are also Nash equilibria, although not all Nash equilibria are subgame perfect. Nash equilibria that are based on noncredible threats are not subgame perfect. Formally, we define a subgame perfect equilibrium as a strategy profile s^* in the overall game (D) such that for any history h , the profile $s^*_{|h}$ is a Nash equilibrium for the subgame $D(h)$. Intuitively this says the equilibrium for any subgame is rational, conditional on equilibrium play in the future. For example, in the centipede game, the strategy profile of A choosing D (down) at stage 1 is rational, conditional on what we know will happen in future subgames. So this is a subgame perfect equilibrium.

Subgame A segment of a larger game.

REPEATED GAMES

The business world is characterized by repeated interactions. In many markets, firms and managers compete against each other for decades. Within firms, managers interact with each other over long periods. So how does a perception of future interaction affect behavior? Managers need to understand how the prospect of a future can change player actions. Again we turn to backward induction to help us understand the implications of repeated play.

We illustrate the strategic effect of repeated play using a stylized example from a class of games known as prisoner's dilemmas. Suppose Allied and Barkley produce an identical product. They also have similar cost structures. Managers at both firms must decide whether to price the product high or low. This situation is shown in Figure 12.13.

The Nash equilibrium of this game is for managers at both firms to sell their products at a low price (and earn \$3 million). Although both realize they are better off if each maintains a high price, they are afraid that managers at the other firm will then drop the price and steal the market. Hence managers at both firms price low. If Allied and Barkley managers compete in only one market for a single

FIGURE 12.13

Pricing as a Prisoner's Dilemma

		Allied's pricing strategies	
		Price high	Price low
Barkley's pricing strategies	Price high	5, 5	1, 20
	Price low	20, 1	3, 3

instance, we expect both of them to keep prices low. After all, that is the rational choice.

But what if they are rivals for a long time? Instead of playing this pricing game once, they play it multiple times. Should we still expect managers at both firms to price low? Each firm can see it is forfeiting \$2 million per period, simply because neither can trust the other to maintain the high price. As Emerson commented, “Distrust is expensive.”

Strategically, the key difference between one-shot games and those that are repeated is the presence of a future. A future introduces behavior not possible in a one-shot world. Trust, reputation, promises, threats, and reciprocity need a future to exist. A future also means that payoffs are no longer relegated to the short term because we now face longer-term implications. A betrayal of trust may create gains in the present, but these may be outweighed by future losses.

Models of repeated games reflect and account for this wider range of feasible behavior. They use the idea of a future to construct norms that let players reach mutually beneficial outcomes. For example, in the preceding situation, these norms help maintain an equilibrium where both firms price high. The risk of one firm undercutting the other is mitigated by the threat of future punishment. Of course such threats must be credible. Let us see how these models work.

The first distinction these models make is whether or not the time horizon is finite. Cooperative behavior is easier to maintain in an infinite horizon game because the future always looms. In finite horizon games, the future grows smaller as we approach the last period. So consider an infinite horizon in the game shown in Figure 12.13. If Allied and Barkley managers cooperate and price high, each receives a payoff of 5 per period. One can defect, price low, and earn 20 for a single period. The other will then price low, and each will receive 3 for the rest of

the game. So the incremental earnings of 15 ($20 - 5$) are lost within 8 periods ($(5 - 3) \times 8$). In fact, in an infinite horizon game, no single-period noncooperative payoff will be larger than the sum of cooperative future payoffs.

The long shadow of a future in an infinite horizon game causes the well-known result called the **folk theorem**. This theorem basically states that any type of behavior can be supported by an equilibrium (as long as the players believe there is a high probability that future interaction will occur). The support for a wide range of behaviors occurs because the future always matters in these games; hence credible threats and promises can alter the current behavior of players. Of course this makes it much harder to accurately predict behavior in games with infinite horizons.

Finite horizon games are fundamentally different because as the game progresses, the future necessarily grows shorter. Because behavior in these games is predicated on the use of credible signals of future behavior, their power diminishes as the future grows shorter. And in the last period, signals hold no power because there is no future (the last period of a repeated game is akin to a single shot); hence the Nash equilibrium is identical to that of a one-shot game. In the pricing game, this means that managers at both firms price low. Without the restraint of credible signals and a future, managers should expect others to act opportunistically.

But wait. If we know managers at both firms will price low in the last period, how should this affect their actions in the prior-to-last period? Let us again use backward induction. Is there any strategy managers at either firm can follow to change the low-price outcome of the last period? No: Regardless of the strategies played, managers at both firms will price low. Their strategic fate is sealed. If this is true, the rational strategy is to price low in this period, too. And just as we saw in the centipede game, similar reasoning extends backward to the first period. So the equilibrium in a repeated version of the game is identical to that of a one-shot game. Managers at both firms should price low.

But what about the looming future and the use of credible commitments? Shouldn't that change behavior? Game theorists had similar thoughts, so they developed the theory to account for these factors. This was a more difficult task relative to the infinite horizon game because they had to account for the final period.

One insight recognized by game theorists is that not everyone is opportunistic. What happens if nonopportunistic individuals are in the population and you do not know (with certainty) whether you are playing with one? How does that change predicted managerial behavior?

Folk theorem This theorem states that any type of behavior can be supported by an equilibrium (as long as the players believe there is a high probability that future interaction will occur).

INCOMPLETE INFORMATION GAMES

This question established a branch of game theory called **incomplete information games (IIG)**. These games loosen the restrictive assumption that all players have the same information. The introduction of incomplete information makes it possible to derive cooperation (price high in Figure 12.13) as an equilibrium

Incomplete information games (IIG) A branch of game theory that loosens the restrictive assumption that all players have the same information.

Tit for tat Strategy that allows players to cooperate in the first period and in all succeeding periods the players mimic the preceding period's strategy of the other player.

behavior. Now, when we backwardly induct, pricing low is not necessarily the predicted strategy in the last period. Nonopportunistic players may still price high in the last period because they obtain satisfaction from cooperating. They do not care that there is no future; they just want to cooperate. Instead of pricing low from the initial period, players may want to experiment in early periods by pricing high to determine whether they are playing with a nonopportunistic type.

One strategy players may use to experiment is commonly referred to as **tit for tat**. Players using tit for tat cooperate in the first period. In all succeeding periods they mimic the preceding period's strategy of the other player. For example, assume Barkley managers use a tit-for-tat strategy. In period 1 they price high. In period 2 Barkley managers mimic Allied's period 1 strategy. In period 3 they mimic Allied's period 2 strategy. So Barkley managers begin the game by pricing high. They continue to price high as long as Allied managers price high. If Allied managers price low, they price low in the following period and continue to price low until Allied managers price high again. Using this strategy, Barkley managers determine whether those at Allied are opportunistic and, if they are, suffer only one period of low payoffs.

In IIG models, players possess asymmetric information. For example, Barkley managers may know more details about their own cost function than Allied managers know. IIG models summarize this asymmetric information in the form of player types. A *type* consists of player characteristics that are unknown to others. In business, types may consist of competitive attributes, like cost functions. In personal relationships, they may consist of personality traits, like trustworthiness.

Specific types are represented by different payoff (preference) functions. So a low-cost type has a different payoff function than a high-cost type. A simple IIG model is shown in Figure 12.14. Allied managers need to decide whether to enter a product market where Barkley is an incumbent. Allied managers are uncertain of the reaction of Barkley managers if they decide to enter the market. If Barkley managers are "tough," those at Allied expect them to lower their price and defend their market. If Barkley managers are "soft," those at Allied expect them to keep their price high, thereby allowing Allied to enter the market.

When Barkley managers are actually tough (part A of Figure 12.14), the Nash equilibrium is for Allied managers not to enter the market and for those at Barkley to price low. Of course Barkley managers will not have to price low because those at Allied will never enter the market. When Barkley managers are actually soft (part B of Figure 12.14), the Nash equilibrium is for Allied to enter the market and for Barkley to price high (allowing entry). Note that Allied payoffs are identical across parts A and B. The incomplete information is about Barkley managers, not those at Allied. Only the payoffs of Barkley managers change.

So if Allied managers know the true type of Barkley managers, their decision is easy. If Barkley managers are soft, they enter the market; if they are tough they don't. Unfortunately, Allied managers aren't sure which type is true. And knowing this, can Barkley managers influence their beliefs?

FIGURE 12.14

Tough or Soft Barkley Managers

A. Barkley managers are tough.

		Allied's strategies	
		Enter the market	Do not enter the market
Barkley's strategies	Fight (price low)	6, 2	8, 3
	No fight (price high)	5, 4	2, 3

B. Barkley managers are soft.

		Allied's strategies	
		Enter the market	Do not enter the market
Barkley's strategies	Fight (price low)	2, 2	3, 3
	No fight (price high)	4, 4	7, 3

REPUTATION BUILDING

The presence of a future and incomplete information are the necessary ingredients for building reputations. In their presence, a reputation is a rent-generating asset. In the example here, if Barkley managers convince Allied managers they are tough in the early periods of the repeated game, Allied managers will stay out of the market in the later stages.

In game theory, a reputation is simply the history of behavior. Intuitively, reputation-building models parallel the human thought process. When we are unsure about the traits of others, we look to past behavior for clues. We use this information to form probabilistic beliefs regarding the traits of others. "I think I can trust him, but I wouldn't bet my life on it." In effect, we use a reputation model to infer future actions from past behavior. For example, assume a friend asks you to lend her \$100. If you have previously lent money to this friend, you

PROBLEM SOLVED: Can Rating Agencies Improve a Bank's Capital Structure?

Suppose banks A and B compete with each other for high-wealth clientele by simultaneously choosing how much capital each should hold. Because of the rich clientele, neither bank is regulated or rated by the industry rating agency Unstandard and Rich (U&R). If both banks choose HIGH capital, then consumers will perceive each to be safe and there will be little systemic risk and both banks will earn profits of 50. If both banks choose LOW capital, each will reduce its cost of capital, but each will be perceived by clients as riskier and each will earn profits of 20. If one bank chooses LOW and the other HIGH, the LOW capital bank will be able to offer more attractive terms and increase its profit to 60, while the HIGH capital bank will earn only 10.

We can use game theory to determine what capital each bank will choose. A's best choices given B's choices are underlined. B's best choices given A's choices are bolded.

		Bank A	
		HIGH capital	LOW capital
Bank B	HIGH capital	50, <u>50</u>	10, <u>60</u>
	LOW capital	60 , 10	20 , <u>20</u>

The Nash equilibrium is both banks choosing LOW capital. It's a prisoner's dilemma game. The win-win scenario is both banks choosing HIGH capital; but this scenario will not occur given the dominant strategies of each bank to choose LOW capital.

Is it possible to change the game in a way that allows the banks to maximize profits, that is, to both choose High capital? Suppose the directors of Bank A realize that, without regulation and rating, clients do not fully realize how risky a bank will be if it carries low capital. With such a LOW rating by U&R, clients would fully realize the true risk and profits would be reduced by 15 from the numbers given above (the profit is 15 lower whatever the strategy chosen by the rival). The profits for HIGH capital are as before. The directors realize that, if they choose to be rated voluntarily, the decision is irreversible and Bank B would be forced to follow. Does the decision by the directors of Bank A to be rated by U&R advantageously change the game for the banks? Indeed, it does, as can be seen in the matrix below where A's best choices given B's choices are underlined. B's best choices given A's choices are bolded.

		Bank A	
		HIGH capital	LOW capital
Bank B	HIGH capital	50, <u>50</u>	10 , 45
	LOW capital	45, <u>10</u>	5, 5

Bank A would choose to be rated. Bank B would follow. With ratings by U&R, each bank has a dominant strategy to choose high capital. Thus, both Bank A and Bank B make a profit of 50.

The lesson learned is that providing information that is valuable to both the banks and their clients can improve payoffs.

will recall whether she paid you back. You are more likely to lend money to a friend who repaid a previous debt. Why? Because the friend paid back the earlier debt, you perceive a higher probability that he or she will pay back this new debt. If the previous debt was not repaid, you are less likely to lend the person money again.

So in situations with futures and incomplete information, managers need to generate reputations to earn future rents. In all such situations, using backward induction, managers need to consider how current behavior will affect the future. Reneging on a debt has immediate payoffs (the debtor gains the amount lent) and long-term consequences (the lender is less likely to offer money in the future). Although these reputation models are too complex to explain here, the underlying idea is simple. In games with a future, players must consider both the present and the future. The payoff managers generate has two components: the immediate gain and its effect on future gains.

For example, suppose Barkley's managers are actually soft. They still have an incentive to act tough in early periods. Of course this will give them a lower payoff in these early periods (3 in red) than playing their true soft type (4 in red). But if they act tough early, Allied managers might not enter later because they are convinced (at least enough) that Barkley managers are truly tough. This allows Barkley managers to earn 7 in later periods. Note, though, that as the future gets shorter (as it necessarily does in finite horizon games), the value of maintaining a false reputation shrinks. So in these later periods, there is an increasing probability that Barkley managers will reveal their true type to be soft. And in the final period, Barkley managers will definitely reveal they are soft.

Examples of managerial use of reputation building are easily found. From product quality to entry deterrence, corporate culture to honest auditors, these models help explain behavior. For example, long before the recent corporate fraud scandals, game theorists modeled auditing firms as renting their reputations for being honest. The models predicted that any accounting firms involved in fraudulent activities would lose their high-quality reputations, and the value of their names would decrease toward zero. This is exactly what we saw with the implosion of Arthur Andersen after the accounting scandal at Enron.

COORDINATION GAMES

It is often profitable for managers to coordinate actions with others. Although this is certainly true for activity within a firm, it is often true for market strategy. Managers must consider the benefits and costs of coordination efforts. Managing this effort is an important managerial task. Managers engage in many coordination games; it is essential that they understand their payoff structure.

Game models represent coordination games as containing more than one Nash equilibrium. Recognizing the Nash equilibria (that is, the outcomes on which managers want to coordinate) is generally not an issue, but choosing which

one to select is. We will see that as game parameters change, the impediments to coordination shift. Game theory visualizes these shifts with changes in the payoff structure.

Matching Games

In matching games, players generally prefer the same outcome. However, there may be impediments to reaching this outcome. Impediments may include the inability to communicate, different ideas about how to reach an objective, or asymmetric information. The game we show in Figure 12.15 concerns coordination of product attributes.

This game has two Nash equilibria. We expect one firm to produce for the consumer market and one for the industry. Though 7, 7 is clearly inferior to 12, 12, it is not ruled out as a Nash equilibrium. However, note that both Allied and Barkley prefer a payoff of 12 to that of 7.

Battle of the Sexes

In this coordination game, players still want to coordinate, but they prefer different outcomes. Because of different preferences, each prefers a payoff not favored by the other. If this game is repeated, players often switch between equilibria so that both gain. However, in one-shot games like Figure 12.16, it is more difficult to predict the outcome.

Similar to the matching game, each wants to enter the submarket not entered by the other, but now the payoffs are not equal. Managers at both Allied and Barkley are better off if they produce a high-end product, so it is not clear on which outcome they will coordinate.

FIGURE 12.15

Product Coordination Game

		Allied's strategies	
		Produce for consumer market	Produce for industrial market
Barkley's strategies	Produce for consumer market	0, 0	7, 7
	Produce for industrial market	12, 12	0, 0

FIGURE 12.16

Battle of the Sexes

		Allied's strategies	
		High-end product	Low-end product
Barkley's strategies	High-end product	0, 0	11, 6
	Low-end product	6, 11	0, 0

Assurance Games

Coordination games like that in Figure 12.17 are also known as stag hunt games. The French philosopher Rousseau tells the story of two hunters (actually poachers) who could coordinate efforts and catch a stag or renege on their agreement and each catch rabbits for himself. Although each prefers to catch the stag, that strategy carries the risk that the other will renege and the first hunter will catch nothing. So players have similar preferences for outcomes but have an associated risk.

Here we model a decision of whether to shift to new standards. Although managers at both firms prefer to shift, there is a risk if one shifts and the other does not. We say the outcome 12,12 is Pareto dominant (both players are better

FIGURE 12.17

Stag Hunt or Assurance Game

		Allied's strategies	
		Stay with old standard	Shift to new standard
Barkley's strategies	Stay with old standard	6, 6	6, 0
	Shift to new standard	0, 6	12, 12

off) but risk dominated (if one chooses to shift and the other does not, the firm shifting receives 0).

First-Mover Games

We can also use coordination games to show the benefits of moving first. Figure 12.18 shows a sequential game in which managers at both firms want to coordinate but each has an incentive to produce a superior product (similar to the battle of the sexes). However, in this game it is possible to move first by speeding up product development. The game shows which firm managers will move first and how much they are willing to pay to speed up the process.

In this game, managers at both firms want to introduce the superior product first. There are two Nash equilibria. Allied managers produce a superior product and those at Barkley produce an inferior one, or vice versa. Once one firm produces a superior product, the other is resigned to producing an inferior product. The question is this: Which firm's managers are willing to pay a higher price to produce the superior product? We can answer this question by looking at the incremental benefits of moving first. The incremental benefit to Allied managers of producing the superior product is the difference in payoff between producing the superior product (\$140) and the inferior product (\$70). This is a difference of \$70. The incremental benefit to Barkley managers of moving first is $\$110 - \30 or \$80. We would predict that Barkley will move first because it is willing to spend up to \$80 to move first, whereas Allied will spend only up to \$70.

Hawks and Doves

This interesting coordination game has been applied to behavior in both human and animal worlds. Assume two players are locked in a conflict. If both players act

FIGURE 12.18

First-Mover Advantage

		Allied's strategies	
		Produce superior product	Produce inferior product
Barkley's strategies	Produce superior product	25, 50	110, 70
	Produce inferior product	30, 140	20, 30

FIGURE 12.19

Hawks and Doves

		Country 1 strategy	
		Act like a hawk	Act like a dove
Country 2 strategy	Act like a hawk	-1, -1	10, 0
	Act like a dove	0, 10	5, 5

like hawks, conflict is inevitable. However, if one acts like a hawk and the other backs down (acting like a dove), conflict is avoided. If both are doves, conflict is not even threatened. The game is shown in Figure 12.19.

There are two Nash equilibria; they require one country to act like a hawk and the other to act like a dove. The issue is which country will back down and act like a dove because this country will suffer a lower payoff.

John Maynard Smith applied similar models to the animal kingdom to model the fighting behavior of animals. One interesting example concerns the behavior of spiders in New Mexico. Webs are a scarce commodity within the spider community because they are difficult to build in the desert. However, a female spider needs a web to lay her eggs. Therefore, female spiders fight (or threaten to fight) over existing webs. They do so by approaching a web and violently shaking it. After each shows this force, one spider (the dove) generally leaves the web to the other spider. Rarely do the spiders actually engage in a physical fight. Smith and other biologists noted that certain physical traits account for which spider is the hawk and which is the dove. For the spiders, the two most important traits appear to be incumbency and weight. The heavier spider usually claims the web, whereas the lighter spider backs down. Smith believes that the violent shaking of the web is actually a credible signal of which spider is the heaviest.

STRICTLY COMPETITIVE GAMES

The games we have just looked at have mixed motives in the sense that conflict interfaces with mutual dependence. However, some games are strictly competitive: Any gain by one player means a loss by another player. The net gain is always zero;

FIGURE 12.20

Advertising Campaigns

		Allied		
		Campaign A	Campaign B	Campaign C
Barkley	Campaign 1	-5, 5	20, -20	-22, 22
	Campaign 2	-3, 3	7, -7	4, -4
	Campaign 3	-4, 4	-6, 6	17, -17

Zero-sum games A competitive game in which any gain by one player means a loss by another player.

what one gains, the other loses. These games are also known as **zero-sum games**. For example, slow-growth (mature) markets are characterized as zero-sum. Because the market size remains fairly constant, any increase in the share of one firm means an identical decrease in the share of another firm. Figure 12.20 shows one such example.

Zero-sum games are still solvable using Nash equilibria. In Figure 12.20 the Nash equilibrium is for Allied managers to use campaign A and Barkley managers to use campaign 2.

SUMMARY

1. Strategic decisions involve interactive payoffs. Because a player's payoff depends on his or her decision and the decisions of others, that player must anticipate the actions of others in formulating an optimal strategy.
2. Game theory is a mathematical framework that can help managers anticipate the actions of others. The theory helps managers represent strategic issues by focusing on the players involved, their feasible strategies, the possible outcomes, and the payoffs associated with those outcomes.
3. In solving games, managers first need to look for dominant strategies. If they exist, managers need not consider the actions of others. Rational players always play their dominant strategy.
4. If dominant strategies do not exist, managers should try to predict the behavior of others using the solution concept of the Nash equilibrium. This concept assumes that all players do the best they can, conditional on all others

doing the best they can. This is the most widely used solution concept in game theory.

5. Managers should use strategic foresight; this is the ability to make decisions today that are rational, conditional on the anticipated future behavior of others. Game theory models this foresight through backward induction. In using backward induction, we go to the end of the game to determine what strategies players will use, then choose an action for the current period that is rational given these future beliefs.

6. Managers must pay attention only to signals that are credible. Game models can determine the credibility of threats, promises, commitments, and the like.

7. Games with a future are called *repeated games*. When a future exists, players may change the strategies they select. Generally speaking, gaining cooperation from others is much easier in a repeated game.

8. Game theorists have developed incomplete information models to look at situations where there is a future and some uncertainty exists about the traits of others. Under these conditions, building a reputation is important because reputations can generate future rents.

9. The ability to coordinate is an important managerial trait. Coordination models help managers better understand the impediments to coordination and the actions necessary to decrease coordination costs.

PROBLEMS

1. Two soap producers, the Fortnum Company and the Maison Company, can stress either newspapers or magazines in their forthcoming advertising campaigns. The payoff matrix is as follows:

		Maison Company	
		Stress newspapers	Stress magazines
Fortnum Company	Stress newspapers	\$8 mm, \$9 mm	\$7 mm, \$8 mm
	Stress magazines	\$9 mm, \$8 mm	\$8 mm, \$7 mm

- a. Is there a dominant strategy for each firm? If so, what is it?
- b. What will be the profit of each firm?
- c. Is this game an example of the prisoner's dilemma?

2. The Ulysses Corporation and the Xenophon Company are the only producers of a sophisticated type of camera. They each can engage in either a high or a low level of advertising in trade journals. The payoff matrix is as follows:

		Xenophon Company	
		Low level	High level
Ulysses Corporation	Low level	\$12 mm, \$13 mm	\$11 mm, \$12 mm
	High level	\$13 mm, \$12 mm	\$12 mm, \$11 mm

- Will Ulysses engage in a high or a low level of advertising in trade journals?
 - Will Xenophon engage in a high or a low level of advertising in trade journals?
 - Is there a dominant strategy for each firm?
3. The *New York Times* reports that Wal-Mart has decided to challenge Netflix and enter the online DVD-by-mail market. Because of economies of scale, Wal-Mart has a slight cost advantage relative to Netflix. Wal-Mart is considering the use of a limit pricing strategy. It can enter the market by matching Netflix on price. If it does, and Netflix maintains its price, then both firms would earn \$5 million. But if Netflix drops its price in response, Wal-Mart would have to follow and would earn \$2 million; Netflix would earn \$3 million. Or Wal-Mart could enter the market with a price that is below Netflix's current price but above its marginal cost. If it does, Netflix would make one of two moves. It could reduce its price to below that of Wal-Mart. If it does, Wal-Mart will earn a profit of \$0, and Netflix will earn a profit of \$2 million. Or Netflix could keep its present price. If Netflix keeps its present price, Wal-Mart can keep its present price and earn \$6 million (while Netflix earns \$4 million). Or Wal-Mart can increase its price and earn \$2 million while Netflix earns \$6 million.
- Draw the extensive form of this game and solve it.
 - Draw the game's matrix form and identify any Nash equilibria.
4. Two rival PE firms are interested in funding the same two start-ups. Each would prefer not to get into a bidding war with the other regarding either of the start-ups. Payoffs are given in the following table:

		PE Firm A	
		Invest Start-up 1	Invest Start-up 2
PE Firm B	Invest Start-up 1	10, 10	60, 40
	Invest Start-up 2	25, 55	20, 20

What are the reservation prices of managers at the two firms? What will bidding look like?

5. Two soft drink producers, York Cola and Reno Cola, secretly collude to fix prices. Each firm must decide whether to abide by the agreement or to cheat on it. The payoff matrix is as follows:

		York Cola	
		Abide by agreement	Cheat
Reno Cola	Abide by agreement	\$29 mm, \$29 mm	\$26 mm, \$30 mm
	Cheat	\$30 mm, \$26 mm	\$28 mm, \$28 mm

- What strategy will each firm choose, and what will be each firm's profit?
 - Does it matter whether this agreement is for one period or for three periods?
 - Is this game an example of the prisoner's dilemma?
6. **Part 1:** Firm A currently monopolizes its market and earns profits of \$10 million. Firm B is a potential entrant that is thinking about entering the market. If B does not enter the market, it earns profits of \$0, while A continues to earn profits of \$10 million. If B enters, then A must choose between accommodating entry and fighting it. If A accommodates, then A earns \$5 million and B earns \$5 million. If A fights, then both firms lose \$5 million. **Draw the game in extensive form and predict the outcome.**

Part 2: Again, consider the above game. Now, suppose the decision of B to enter is reversible in the following way. After B enters the market, and A has decided to either fight or accommodate, B can choose to remain in the market or exit. All payoffs from the above game remain the same. However, if B decides to exit the market, then B suffers a loss of \$1 million, while A regains its old profits of \$10 million. **Draw the game in extensive form and predict the outcome.**

7. The Rose Corporation is one of two sellers of paint. It pursues a tit-for-tat strategy. However, it has great difficulty in telling whether its rival is secretly cutting its price. What problems is this likely to cause?
8. Consider a father who is trying to discipline his child. The father insists that the child must go with the rest of the family to visit their grandmother. The child prefers to go to the movies with a friend. The father threatens to punish the child if the child doesn't visit the grandmother. If the child goes with the family to visit the grandmother, both the child and the father receive one unit of utility. If the child refuses to go to the grandmother's house, and the father punishes the child, the child receives one unit of utility, and the father receives one unit of utility. If the child refuses to go and the father relents (does not punish), the child receives two units of utility, and the father receives none.
 - a. Draw this game in matrix form.
 - b. Draw this game in extensive form.
 - c. Solve this game via backward induction.
9. The Boca Raton Company announces that if it reduces its price subsequent to a purchase, the early customer will get a rebate so that he or she will pay no more than those buying after the price reduction.
 - a. If the Boca Raton Company has only one rival, and if its rival too makes such an announcement, does this change the payoff matrix? If so, in what way?
 - b. Do such announcements tend to discourage price cutting? Why or why not?

EXCEL EXERCISE: GAME THEORY

After airline deregulation in the late 1970s, new entrants would enter or consider entering the market. Most new entrants failed, but Southwest, JetBlue, and AirTran (now merged with Southwest) are among the few exceptions. The legacy carriers (those in the market before deregulation) took particular affront when a new entrant attempted to enter or threatened to enter a hub airport of the legacy carrier. Traditionally, the new entrants were undercapitalized and could not tolerate a significant period of losses.

Several carriers entered American Airlines' hub at Dallas-Fort Worth. American had to consider whether to accommodate the entrant or whether to fight the entrant. Suppose the matrix below represents one time period of a potentially multiple-period game.

		Potential Entrant	
		Enter	Don't Enter
American	Accommodate	2, 1	3, 0
	Fight	1, -1	2, 0

Following the analysis in the text, we see that American has a dominant strategy to accommodate. The potential entrant does not have a dominant strategy, but when the entrant puts itself in American's shoes, the potential entrant determines that American has a dominant strategy to let it enter the market. Given that American will let it in, it'll enter (since $1 > 0$).

Did American accommodate entry into Dallas-Fort Worth? No. Because it was a multiple-period game, and American knew these potential entrants could not tolerate many periods of losses (because they were undercapitalized—they did not have “deep pockets”), American fought them. Under the fight scenario, the potential entrant loses -1 each period it is in the market. Suppose the potential entrant can take two periods of losses and then must fold.

If American accommodates the potential entrant, American's profit stream over time is

2, 2, 2, 2, 2, 2, 2,

If American fights the potential entrant, American's profit stream over time is

1, 1, 3, 3, 3, 3,

So the question is whether the net present value of the second stream is better than the net present value of the first stream, that is,

1, 1, 3, 3, 3, 3, $>$ 2, 2, 2, 2, 2, 2,

Netting out terms, this becomes

0, 0, 1, 1, 1, 1, $>$ 1, 1, 0, 0, 0, 0,

The discrete discount factor to determine the net present value of cash in period t today is $1/(1 + i)^t$, where i is the discount (think interest) rate. We'll assume the first period in the analysis above is time 0. Thus, the net present value of the question asked above is

$$\begin{aligned}
 &0 + 0*[1/(1 + i)^1] + 1*[1/(1 + i)^2] + 1*[1/(1 + i)^3] \\
 &+ 1*[1/(1 + i)^4] + \dots > 1 + 1*[1/(1 + i)^1] + 0*[1/(1 + i)^2] \\
 &+ 0*[1/(1 + i)^3] + 0*[1/(1 + i)^4] + \dots
 \end{aligned}$$

or

$$1*[1/(1+i)^2] + 1*[1/(1+i)^3] + 1*[1/(1+i)^4] + \dots > 1 + 1*[1/(1+i)^1] + 0 + \dots$$

or

$$-1 - 1*[1/(1+i)^1] + 1*[1/(1+i)^2] + 1*[1/(1+i)^3] + 1*[1/(1+i)^4] + \dots > 0$$

So what discount rate makes the above equation positive?

This is a job for a spreadsheet. We'll try this out for five periods.

Enter $=-1-1*(1/(1+A2)^1)+1*(1/(1+A2)^2)+1*(1/(1+A2)^3)+1*(1/(1+A2)^4)$ in cell A1. Then play around with various interest rates until in cell A2 cell A1 reads 0.

If the above periods are years, a yearly interest rate of approximately 17.87% per year will make the equation in cell A1 zero. Internal rates of return for firms in this range are not unusual, so any rate higher than that would make fighting entry a good long-run strategy. We could play the game for fewer periods (which would lead to a higher interest rate answer) or for more years (which would lead to a lower interest rate answer). We could also play with different payoff numbers for American in the above game matrix. The point is that many different what-if scenarios can be calculated quite quickly by using a spreadsheet.

The other benefit of the fight policy is that American gains a reputation that it will fight any entrant. If this reputation sticks, American will not be faced with the 1, 1, 3, 3, 3, profit stream scenario but won't have to fight and thus will face a profit stream of 3, 3, 3, 3, 3,

AUCTIONS

To illustrate the worth of game theory to managerial thinking, let us use it to examine behavior in auctions. All auctions, like markets, are governed by rules and procedures. Managers who understand these rules use them to create strategic advantage; for instance, they design auctions to generate higher profit or they bid more efficiently to capture greater surplus. We focus here on the auction mechanism because this parameter significantly influences behavior in auctions. We also show that managers need not abandon our sophisticated pricing strategies in auction settings. Performance-enhancing strategies like price discrimination are inherent in many auction designs. For the manager, it is a matter of understanding auction design and incentives.

It goes without saying that managers in financial services must understand auction design because this is the heart of financial markets. The growth of auctionlike mechanisms in markets outside the financial realm makes it imperative that all managers understand auction design. The use of auctions in the United States has grown roughly 10% a year since 2005. Auction revenue in 2008 totaled over \$278 billion, with roughly \$10.2 billion generated from online auctions.¹ Managers will also find an understanding of auctions helpful in other areas. We will use the theory to develop behavioral implications for negotiations and monopoly markets.

LEARNING OBJECTIVES

- A Short History of Auctions
- Types of Auction Mechanisms
- Auction Mechanism and Revenue Generation
- Bidding Strategies
- Strategies for Sellers
- Value of Information
- Risk Aversion
- Number of Bidders
- Winner's Curse
- Concerns in Auction Design
- Summary
- Problems
- Excel Exercise: Auctions

1. Auction Industry 2010, Forrest Carter, Broad College of Business Report, 2010.

A SHORT HISTORY OF AUCTIONS

The first written record of an auction was the annual Babylonian marriage market described by the Greek historian Herodotus. The socially aware Babylonians structured the auction to ensure that all women who wanted husbands found them. The most beautiful women were auctioned first, their bidders being the rich who could afford to pay for beauty. Less attractive women then participated in a negative bid auction. The man with the lowest negative bid received the woman at his bid price. This price was then paid from funds generated by the beautiful women. Not only did this ensure that all willing women were married; it also gave the newlyweds financial support.

Auctions were used in ancient Greece to award mineral rights and by Roman authorities to collect debts owed by individuals. Cicero reports a court case in 80 BC involving the auctioning of goods to satisfy a debt. In 1556 the French monarch appointed an officer to appraise and auction all property left by those executed by the state. Pepys reports on a candle auction of 1660. Candle auctions gave bidders a limited time to bid (until the candle burned out); they were the forerunner of today's timed auctions on eBay.

Auctions almost caused a continuation of the War of 1812. During the war, English goods were prohibited from being imported to the United States. The end of the war saw British merchants flooding the United States with goods to satisfy pent-up demand. The goods were sold via auction because of the speed at which goods can be allocated. American manufacturers claimed the British were dumping the goods (selling them below cost) and tried to get the auctions outlawed. Newspapers and U.S. trade organizations led a spirited fight to have state legislatures abolish auctions. The fight extended to the U.S. Congress, where the Ways and Means Committee introduced a bill to ban auctions. Fortunately Congress never passed it.

Although markets grew more quickly than auctions, auctions played a major role in allocating fruits and vegetables, fish, furs, tobacco, and livestock. By the 1980s auctions were used to sell over \$5 billion of goods yearly.

Auction use continued to increase and virtually exploded in the late 1990s with the increase in e-commerce. The Internet is to auctions what mass retailing is to posted prices: a highly productive marriage of technology and allocation mechanism. Internet auction use can significantly decrease organizational costs. In most auctions, if a manager is the seller, the more buyers the merrier (this idea is shown formally later in the chapter). Expected revenue does not decrease with increases in the number of bidders. So any technology that decreases the costs of bringing individuals together should increase auction use.

STRATEGY SESSION: The Expanding Use of Auctions

The first baseball game of the season—the home opener—is usually one of the “hot ticket” games of the season. Tickets are in high demand because many people who aren’t even serious fans want to be seen at a happening event. The New York Mets team came up with a novel way of selling 100 of the 42,000 first game tickets in their brand new Citi Field in 2009. They held an online auction. A bidder could bid for two or four tickets. The Mets usually have a tiered pricing structure for these 100 seats at \$525, \$315, \$280, and \$210. The minimum bids received by midday before the game were \$610, \$460, \$350, and \$310.

The New York Jets football team has a new stadium (shared with the New York Giants) in the New Jersey Meadowlands. Both teams used personal seat licenses (PSLs—two-part tariffs) to help finance the stadium and to determine who had the rights to purchase tickets for the games. The question was how to distribute the licenses and at what price to sell them. The Jets sold some of their licenses via an online auction on StubHub (an eBay subsidiary). In a nine-day auction in late October 2008 (in a very depressed economy in New York), 620 of the 2,000 available licenses transacted. The average price was \$26,000, with the top price \$82,000 and the lowest price \$10,500. The Jets sold other PSLs at a flat rate ranging from \$4,000 to \$25,000. The auction seats were considered more prime than the \$25,000 seats. In a high-profile kickoff of the online auction, the Jets auctioned off two 50-yard-line prime seats for \$200,000 apiece. Although the objective was to auction off all 2,000 seats, given the state of the economy, both the Jets and StubHub expressed sat-

isfaction with the results. The Giants sold their PSLs at fixed prices (the top being \$20,000).

A large Panamax ship (one that just fits in the current Panama Canal) can cost more than \$40,000 per day to operate. When demand for passage is high, queues of ships can form lasting up to a week. Container cargo of valuable consumer goods has high inventory carrying costs. Sometimes the combination of high ship operating costs and high carrying costs on goods waiting to be delivered can be such that a ship desires to “jump the queue” and become first in line. So far the record paid to jump the queue is \$331,200, paid by a cruise ship the *Disney Magic* (owned by the Walt Disney Company). Apparently, the most valuable cargo is people. The previous record was held by the container ship *Norwegian Pearl*, at over \$313,000. These payments are above and beyond the normal canal toll, which is based on weight. The right to go first can be purchased, as we discussed in Chapter 11, which dealt with the first-mover advantage.

Sources: “Met Auctioning Tickets to Citi Field Opener,” at <http://msn.foxsports.com/mlb/story/9424844/>; Darren Rovell, “NJ Jets Scale Back on PSL Auction: Sign of Economic Times?” at www.cnn.com/id/27363426/; “Jets Earn Over \$16 Million in OnLine PSL Auction,” at <http://nbcports.msnbc.com/id/27421535/>; “US Cruise Ship Pays Record to Cross Panama Canal,” at <http://panamasol.com/us-cruise-ship-pays-record-to-cross-panama-canal/454/>.

TYPES OF AUCTION MECHANISMS

All auctions involve a bidding process. In most auctions, a service or product is made available, and buyers bid for the ownership rights. Reverse auctions occur when a buyer announces the need for a product or service and sellers bid for the right to sell the buyer the good or service (and the low bidder wins the auction). Whether using a normal or reverse auction, managers primarily use one of four auction mechanisms: the English or ascending-bid auction; the Dutch or descending-bid auction; the sealed-bid auction; and the second-price, sealed-bid auction.

English or Ascending-Bid Auction

In English auctions, the initial price is set at the seller's reservation price (called the **reserve price**). Buyers then bid against each other with a succession of higher prices until one bidder remains. This final bidder receives the good or service at his or her final stated price. There are several ways to manage this bidding procedure. An auctioneer can orally call out bids; bidders signal their acceptance of the price with a shake of their heads or a wave of their bidding cards. This method is used often by large auction houses like Christie's and Sotheby's. Another method is for bidders themselves to call out their bids. This method is used in many commodity exchanges. Individual commodities are assigned their own area, called a pit; traders of a commodity gather in the pit and call out the price at which they are willing to buy (sell) the commodity. If others wish to sell (buy) at that price, they do so. However, these pits are increasingly being replaced by more efficient electronic auctions. *The New York Times* reports that the percentage of pit trades in livestock commodities declined from roughly 90% of total trades in 2000 to 22% in 2007.² During this time, electronic markets hastened the closing of commodity pits in Chicago, Hong Kong, London, Sydney, and Tokyo. Exchange operators find that digital auctions have lower operating costs and generate more revenue.

Two special types of the ascending-bid auction are the Japanese auction and the ascending-bid timed auction. Much theoretical work has centered on the **Japanese auction**, in which bidders bid until the price exceeds their reservation price. They then drop out of the bidding process and do not return. Remaining bidders can thus determine who has dropped out and who remains. In the **ascending-bid timed auction**, the bidding continues for a specified time. The high bidder at the end of this period receives the good or service (assuming the bid is greater than the seller's reserve price). These timed auctions are used extensively by Internet auction sites.

A nice example of how managers influence behavior through auction design is shown in behavior at the ascending-bid auction sites of eBay and uBid. eBay managers have chosen to end auctions at preannounced times. This choice encourages

Reserve price The lowest price at which a seller is willing to sell a product; also called a reservation price.

Japanese auction Auction in which bidders bid until the price exceeds their reservation price.

Ascending-bid timed auction Auction in which the bidding continues for a specified time.

2. Niko Koppel, "In Chicago, a Rowdy Trading Scene Grows Quieter," *New York Times*, October 29, 2007.

the practicing of **sniping**: Bidders use programs to ensure that they submit a last-second best bid. Many times this is their only bid. uBid managers also announce a predetermined ending time. However, they use a different rule to end the auction: Auctions at uBid continue as long as there are new bids. Sniping merely extends the time the market is open; the uBid rules blunt its effectiveness as a strategy.

Sniping When bidders use programs to ensure that they submit a last-second best bid.

Dutch or Descending-Bid Auction

In Dutch auctions, initial prices are set very high. A price is announced and a set time passes (for example, 15 seconds). If no bidder accepts the good at that price, the price is lowered by a set interval (for example, a euro) and the procedure repeats itself. The price is lowered until one bidder accepts the announced price. A descending-bid system is often called a **Dutch auction** because it is used to sell flowers in the Netherlands, where auction halls larger than 10 football fields are used daily to auction thousands of flowers. Current offered prices are shown on a large screen, and bidders accept the prices by pushing electronic buttons. The process is very quick with over 500 transactions every hour.

Dutch auction A descending-bid system in which initial prices are set very high and are lowered by set intervals until accepted by bidders.

Sealed-Bid Auction

In these auctions, bidders submit price bids known only to themselves. So unlike an English auction, bidders do not know the valuations of others. In a first-price, sealed-bid auction, bids are opened at a preannounced time, and the highest bidder receives the item at his or her stated price. Reverse sealed-bid auctions are often used to procure goods and services. A state agency may solicit bids for highway construction, with the winning bidder being the one with the lowest bid.

Second-Price, Sealed-Bid Auction

This auction is a variant of the sealed-bid type. In a second-price auction, the highest bidder receives the good or service at the bid price of the second-highest bidder. These auctions are also called **Vickrey auctions** after William Vickrey, who wrote a seminal paper on the subject in the 1960s, which later led to his winning the 1996 Nobel Prize in economics. As we will see, second-price auctions have characteristics conducive to bidders truthfully revealing their valuations.

Vickrey auctions Second-price, sealed-bid auctions in which the highest bidder received the good or service at the bid price of the second-highest bidder.

AUCTION MECHANISM AND REVENUE GENERATION

Similar to market settings, in auctions managers want to maximize profit. And just as differences in structural parameters like differentiation, entry barriers, and location explain variance in market power (and profitability), auction rules explain variance in generated revenues. An advantage of auctions is that relative to markets, it is easier (and less costly) for managers to influence structural parameters through auction rules. Hence it pays for managers to recognize how auction design influences revenues. We examine the effect of auction rules on revenue

by first looking at a simplified model. We then relax some model assumptions to show how rules affect expected revenue.

Our baseline case assumes that bidders are symmetric and risk-neutral, and bids are based on independent signals drawn from a commonly known distribution. The behavioral implications of these assumptions follow:

- **Bidders are symmetric.** Think of a bidder as selecting a valuation (bid) from a distribution of possible bids. Symmetric bidders use similar distributions, commonly known to all. So bidders with identical reservation prices and observing the same signal submit equivalent bids. We emphasize that (as in the real world) differences in reservation prices across bidders are not constrained by this symmetry.
- **Bidders are risk-neutral.** That is, bidders bid to maximize expected values and not risk-adjusted utility. Most managers assume that individuals are risk-averse; however, evidence shows that when individuals decide as corporate agents, behavior can resemble that of risk neutrality. That said, we later show how risk-averse bidders can be strategically exploited.
- **Signals are independent.** Signals are independent in one of two ways. In **private-value auctions**, reservation prices are a function of information and utility. Because signals depend on the information space, valuation is constrained to one's own signal. Say a manager is standing next to another at an auction. Are differences in how the two value a painting a function of their own personal experience, or does what one knows (but is unknown by the other) affect the latter's valuation? The former condition sounds more plausible. This describes a private-value auction with independent signals.

Private-value auctions Auctions in which reservation prices are a function of information and utility.

Common-value auctions Auctions in which all bidders value the good similarly.

In **common-value auctions**, all bidders value the good similarly. What is not known, though, is the true value of the good for which they bid. Consider the rights for minerals under a piece of land. Whatever is down there is worth about the same for all (given worldwide commodity markets), but no one knows perfectly what is there. So each bidder measures and forms beliefs about the true value. If signals are independent, then what one manager estimates does not depend on the estimates of others, although the distribution of signals is commonly known.

For any auction format, let b = bid and p = price paid by the auction winner. Expected profit is simply $(b - p)(Pr_w)$, with Pr_w equal to the probability of winning the auction, conditioned on the bid level. Because bidders want to maximize profit, optimal bids depend on valuations (reservation prices). Therefore, the expected profit for any b, Pr_w combination is given by

$$U(Pr_w, b, p) = (Pr_w)(b - p)$$

Compare these profit functions across auction formats to determine how surplus $(b - p)$ is split between buyer and seller. We later show the slope of the profit

function equals the conditional probability of winning for any given bid. So auction formats with identical surplus functions (conditioned on valuation) offer the same probability of winning at any given bid and hence recommend the same optimal strategies.

This relationship is the foundation for what is called the **revenue equivalence theorem**. The theorem shows that whether a manager chooses an English, Dutch, sealed-bid, or second-price, sealed-bid auction, the choice does not affect the auction's expected total surplus and hence does not affect the expected revenues. The theorem even extends to other auction formats. As long as the format ensures an efficient allocation of goods and gives zero profit to any bidder holding zero value (\$0) for the good, the surplus functions are identical, and so are the recommended bidding strategies across formats. For example, lobbying efforts can be modeled as an "all-pay" auction. All bidders pay for the good, but only the highest bidder receives it. Because this auction satisfies the stated conditions, its expected surplus is the same as any of the four standard auction formats.

Revenue equivalence theorem

Theorem showing that whether a manager chooses an English, Dutch, sealed-bid, or second-price, sealed-bid auction, the choice does not affect the auction's expected total surplus and hence does not affect the expected revenues.

BIDDING STRATEGIES

What are the optimal bidding strategies across the four auction formats? Of the four standard formats, only English auctions let bidders learn more about the reservation prices of others. In the sealed-bid format, bids of others are not revealed until after the bidding is closed. In Dutch auctions, once a bidder reveals his or her reservation price (by accepting the current price), the auction is over. But because of the ascending nature of the English auction and the public nature of bidding, bidders learn more about the reservation prices of others. Unfortunately this information is of limited value to the strategic bidder because optimal behavior is defined by a dominant strategy. And remember, good strategists do not worry about the behavior of others in using a dominant strategy because it lacks strategic value. Thus the optimal behavior in an ascending auction never changes: Managers should always be willing to bid up to their reservation prices.

If used, this guideline maximizes profit ($b - p$). Clearly managers should never bid above their reservation prices. The maximum profit from this strategy is zero, and managers will suffer a negative profit if theirs is the winning bid. Managers do not want to bid higher than necessary because profit decreases with a bid increase. The winner in an ascending auction has to pay only the slightest bit (that is, ϵ) higher than the reservation price of the second-highest bidder. Bids significantly higher than this are inefficient (and suggest badly trained managers). The difference between the reservation prices of the top and second-highest bidders defines available surplus and therefore profit. If managers want to maximize profit, they must capture all available surplus.

The auction's dominant strategy helps managers reduce the complexity of strategic thought. Consider this strategy as a decision rule: Managers listen to any

bid. If the bid is lower than their reservation price, they bid incrementally higher. If the bid is higher, they do not bid. The simplicity of the procedure highlights one component of a dominant strategy's value: process efficiencies. Managers need not consider what others might do in choosing an optimal strategy. So managers can focus (and simplify) the strategic effort, trusting the completeness of dominance.

In any ascending-price auction managers should bid to their reservation prices. If all follow this strategy, the good is sold at the reservation price of the second-highest bidder (actually at an ϵ higher). It is sold to the highest bidder, and his or her expected revenue is the total available surplus, defined as the reservation price differential between the two highest-valuation bidders.

So in an ascending auction, the highest bidder claims the item at the reservation price of the second-highest bidder. We previously noted that this is the same prediction of Vickrey auctions. Thus both English and second-price, sealed-bid auctions are ruled by the same dominant strategy. Though the two auctions differ in process, game theorists predict similar results. Of course the former format is more transparent than the latter.

We now see how the difference in bidding rules affects behavior. In English auctions, the rules allow managers to publicly bid multiple times; in a sealed-bid auction, they bid once privately. English auctions let managers respond to the bids of others; sealed-bid auctions do not. Because managers in a sealed-bid auction have only one chance to bid, they should bid their reservation prices. In both auctions, negative payoffs are nonexistent because zero becomes the manager's worst possible payoff. And the bidder who wins is guaranteed all available surplus. So while the dominant strategy is the same across formats (managers should bid to their reservation prices), the bidding rules of the second-price, sealed-bid auction require different behavior. When auctions have dominant strategies, managers should not worry about others (it achieves nothing); they should focus on their own preferences, think hard about their reservation prices, and then bid them.

Second-price, sealed-bid auctions are a good example of how auction rules influence bidding behavior. The rules in these auctions encourage all bidders to tell the truth. In fact, managers are better off if they tell the truth. In a later discussion we call such rules **incentive-compatible**. This type of rule encourages managers to reveal their true preferences.

Incentive-compatible Type of rule that encourages managers to reveal their true preferences.

Next consider the descending (Dutch) and first-price, sealed-bid auctions. Although the two operate with different bidding rules, they are strategically similar. In fact, consider them identical twins that behave differently. If we model them in their reduced normal or matrix form, they have identical strategy sets and payoffs. That is, any given bid yields the same payoff in either auction, as a function of the bids of others. And unlike the ascending or second-price formats, neither has a dominant strategy. In the Dutch or first-price, sealed-bid auctions, managers must consider what others will bid. Let us see how this design affects a manager's bid choice.

As explained earlier, bidders learn little about the valuations of others in either of these first-price auctions because bids are private until the auction is completed. But because bidders are symmetric, all know the distribution of valuations. This information lets managers estimate the bidding strategy of others. Note that it does not give them the ability to predict the bids of others—just what others would bid conditional on some privately known information.

Each bidder faces an identical decision. If a bid is not the highest, managers receive and pay nothing, so the surplus is zero. If they bid their reservation price and that bid is the highest, the surplus is still zero. A bid higher than their reservation price results in zero surplus (if the manager is lucky and loses the auction) or negative surplus if the manager wins. Managers believe there is some positive probability that their bids are the highest. So the question is this: How should managers structure their bids?

Unlike the previous auctions with dominant strategy, the set of possible strategies approaches infinity in the descending and first-price, sealed-bid auctions. Managers now must consider what others will bid. Bidding strategies become conditional on the bidding strategies of others. The world becomes more complex.

So how do managers reduce the complexity? If all bidders face the same complexity, can bidders help themselves via some constraint on bidding behavior? Economists believe they can by being rational. Rational players consider only those outcomes in which each player maximizes payoffs, given the actions of others. Remember that auction payoffs depend on the actions of all. And unless all managers are choosing their best strategies, given the actions of others, why would any one manager agree to do this? If we adopt this rule that all managers are rational, no individual has an incentive to cheat by changing strategy.

As explained in Chapter 12, in Nash equilibria, all correctly anticipate the actions of others and choose actions that maximize expected surplus. This reduces feasible strategies to a handful and allows managers to focus on fewer scenarios. We note that the Nash concept does not guarantee high payoffs for all. It simply states that given the attributes of an individual and those of others, each individual is maximizing his or her payoff.

So how should managers bid in these auctions according to the Nash equilibrium? Managers clearly should not bid their reservation prices. If they do, their highest possible payoff is zero. Nash assumes that managers will think about their reservation prices and discount backward. The discount should approximate their expectations of the reservation price of the second-highest bidder. Each manager will then bid an ε higher than this belief. Then, if her bid is the highest, she receives positive surplus; if it is not, she receives nothing. As in the ascending and second-price auctions, managers must consider their own reservation prices; but they also must consider what others will bid and base their bids on this expectation. So the decision rule for managers in these auctions is to estimate the reservation price of the second-highest bidder and bid ε higher.

Managerial beliefs about the reservation prices of second-highest bidders are influenced by many factors. One significant factor that should affect beliefs is the number of bidders. As the number of bidders increases, managers need to bid closer to their reservation prices. That is, increased competition reduces expected surplus in these auctions. We can be even more precise if we assume the distribution of bids is uniform. Then the optimal bidding strategy for managers is

$$b = v - \left[\left(\frac{v - L}{n} \right) \right]$$

where v is the bidder's reservation price, L is the lowest possible bid, and n is the number of bidders. For example, assume our own valuation of an item is \$3, and we believe the bids will be evenly distributed between \$0 and \$15. Then our optimal bid if there is only one other bidder is

$$b = 3 - \left[\left(\frac{3 - 0}{2} \right) \right] = \$1.50$$

If there are two other bidders, our optimal bid is

$$b = 3 - \left[\left(\frac{3 - 0}{3} \right) \right] = \$2.00$$

And so we end where we began. In theory, the auction design should not affect expected revenue. The optimal response of bidders in each of the four designs produces the same expected revenue. Managers should bid up to their reservation prices in ascending and second-price auctions because the rules say that if they are the high bidder, they pay only the reservation price of the second-highest bidder. In descending and first-price, sealed auctions, managers consider their reservation prices and then discount back to what they believe are the reservation prices of the second-highest bidders.

STRATEGIES FOR SELLERS

Now we turn to sellers. These days many managers use auctions to sell goods, services, and assets. In these circumstances, good managerial decisions follow the standard economic rule: Maximize profit with marginal revenue equal to marginal cost. Of course some adjustments for the bidding rules must be made.

In markets, managers want to produce where $MR = MC$. This is a quantity decision that determines the pricing point. But auctions are used to sell items few in quantity (often one). The notion of quantity holds little strategic value when quantity is so limited. What determines the optimal pricing point is the distribution of reservation prices across bidders. Therefore managers want to focus on this distribution.

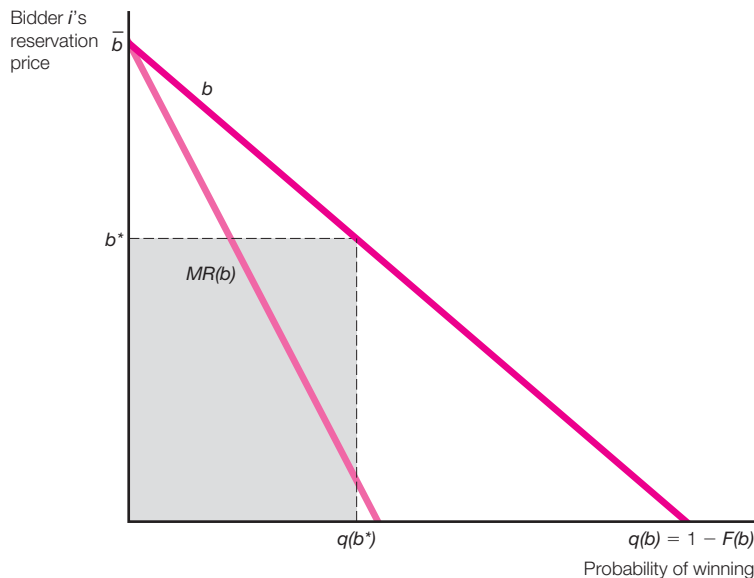
This is a subtle but important difference realized by the strategic manager. Consider an ascending auction. Managerial action in this format is similar to third-degree price discrimination in markets. Recall that efficient allocation is defined as the equality of marginal revenues across markets (the infamous $P_1/P_2 = [1 + (1/\eta_2)]/[1 + (1/\eta_1)]$). That is, managers maximize profit by ensuring the sale to the consumers who value the product or service the most. Total output is determined by the horizontal summation of the marginal revenue curves and its intersection with marginal cost (see Figure 13.1).

Sellers want to use a similar strategy in auctions. Consider an auction selling a unique item. The manager maximizes profit by selling to the bidder with the highest reservation price. Because the seller's marginal costs remain constant for any change in price, the higher the winning bid, the greater the surplus. The expected revenue generated by a given price is simply that price times the probability that it is the winning bid. This, of course, is determined by the distribution of reservation prices across bidders. Consider bidder i and the distribution $F(b)$. Given $F(b)$, we determine the probability that any bid b is the winning bid. This is

FIGURE 13.1

Relationship between the Seller's Expected Revenue and the Winning Bidder's Expected Marginal Revenue

If b^* is the winning bid, the seller's expected revenue is equal to the shaded box. This is also equal to the expected marginal revenue of the bidder.



STRATEGY SESSION: More Managers Use Reverse Auctions to Reduce Costs

We first discussed the usefulness of reverse auctions in 2004 using an online auction site to illustrate the mechanisms. We are happy to report growing legions of managers use this pricing strategy to drive down costs. Managerial use is now spreading from large multinationals to medium-sized firms. And, the spread of broadband has extended the auction's use to consumer markets.

A Gartner analyst counted over 50,000 reverse auctions in the year ending October 31, 2011. This was an increase of over 10% relative to the previous year. Managers are discovering that when used correctly, they help drive down purchase costs, often by up to 20%. Joseph Jimenez, the CEO of Novartis AG, uses them to drive down his purchasing costs for goods and services. There is also evidence that their use in medium-sized companies can generate significant cost savings. For example, Anna's Linens is a chain of 300 stores in the United States. Their managers first used reverse auctions in 2008, and three years later report cost savings of \$6 million.

Managers report benefits other than lower costs. They cite the transparency of markets, the value of better knowing pricing ranges, access to more qualified suppliers, and lower transaction costs. Reverse auctions are best used with commodity-like goods and services.

It is critical that managers think carefully how to specify the good or service needed prior to the auction. Attributes to consider include delivery time, quality, quantity, size, etc. Alan Gladstone, CEO of Anna's Linens, provides more than specifications to potential bidders. After some winning bidders failed to supply the company within its requested delivery time, Gladstone says, "we've added details to the process, not just specifications." Such details include capital and the level of secondary inventory.

There are costs to reverse auctions. There is little relationship building in the auctions; all interaction is transactional. Some bidders miscalculate and find they cannot deliver at the price they bid; though this can be mitigated with good due dili-

simply $1 - F(b)$, as shown in Figure 13.1. Function b acts as a demand curve in the following way. Each point on b suggests the probability of that point being the winning bid. So if the seller were to offer the item at price b , we could predict the expected demand at b . Expected revenue then would be $b[1 - F(b)]$ or the area under the marginal revenue function for any b . Let b^* be the actual winning bid (the reservation price of the second-highest bidder). Mathematically, we can show that b^* represents the expected revenue to the seller and the expected marginal revenue to the winning bidder. That is, in ascending-bid auctions, they are equal.

What does this mean to the seller? As in third-degree price discrimination markets, managers need to sell to high-reservation-price bidders (buyers). And because they want to maximize surplus in both situations, they definitely cannot sell at any price below marginal cost. In auctions, consider marginal cost as the value at which the seller refuses to sell. This is the price at which the seller says, "If that's all others are willing to pay, I'll keep it myself." So sellers need to set a reserve price (or begin the bidding at their reserve price) for the auction. This reserve

gence. There are also the incremental costs of holding an auction.

The use of broadband has decreased the incremental costs of inviting additional suppliers to join the auction. This has allowed smaller firms to join “reverse auction networks” where managers may conduct a reverse auction on a hosted site, where firm members pay a percentage fee to participate. In some countries, consumers have formed their own reverse auction networks to induce tradespeople to bid for jobs. For example, Shiplly.com conducts reverse auctions for delivery services across Europe. Their network consists of over 35,000 couriers and delivery services. Consumers describe what, when, and how they want a package delivered and members of the network bid for the job. Supposedly, this service cuts delivery rates to the consumer by over 50%. The site is very popular with parents who must transport belongings between home and college for their student children. A similar site is called Whocanfixmycar.com. Here consumers offer work to auto mechanics.

After the work is completed by the winning bidder, consumers rate the quality of the service.

Reverse auctions are being used by India to build solar-power farms. On the auction held on December 2, 2011, the lowest bid was for \$147 per megawatt-hour. This bid was 38% lower than the lowest bid in the December 2010 auction. It is estimated the global average price for a solar-power farm is \$208 per megawatt hour. Prices are being driven down due to large overcapacity of photovoltaic cells and equipment. Anand Mahindra, a managing director of a solar company states, “Manufacturers are keen to offer discounts and to defer payments because things are so competitive right now. Costs are going down and that’s really reflected in these bids.”

Sources: Natalie Pearson, “India’s ‘Astonishing Auction’ Pushes Down Global Solar Price,” Bloomberg News, December 2, 2011; Constance Gustke, “Reaping Big Cost Savings with Reverse Auctions,” www.cnbc.com, November 15, 2011; Eric Von Schaper, “Novartis’s Jimenez Has Blockbuster Plans for Diovan After Patent Expires,” Bloomberg News, August 5, 2011.

price should be the bid at which the bidder’s marginal revenue is equal to the seller’s marginal cost. (Recall that the bidder’s marginal revenue is also equal to the auction’s expected revenue.) Hence, under general conditions, managers should consider the following decision rule in setting a reserve price

Optimal reserve price = The value of the object being auctioned off, if it does not sell + (Managerial estimates of the highest reservation price divided by 2)

For example, if the object’s value is zero if it is not sold, and a manager believes the highest reservation price is \$300, he or she should set a reserve price of \$150. After that, sellers should sit back and relax. Let the ascending-bid mechanism work to maximize the expected surplus. Unlike monopoly pricing, but similar to third-degree price discrimination, auctions are often more efficient than a posted-price scheme. In contrast to a posted-price scheme, auctions guarantee that those with the highest reservation price will purchase the good. Note that the managerial similarity to third-degree price discrimination is not confined to ascending

auctions. All formats that are revenue equivalent to ascending auctions share this feature. The following example illustrates the efficiency of auctions.

Assume a seller has four units of output at a marginal cost of \$0. The market has six consumers with reservation prices of \$90, \$60, \$50, \$40, \$20, and \$15. Table 13.1 shows the total available surplus broken down into consumer and seller surplus if the seller uses an auction. Table 13.2 shows the total surplus if the seller posts a \$40 price.

At any posted price other than \$40, the total available surplus decreases. And while the available surplus remains constant with a price of \$40, relative to an auc-

TABLE 13.1

Auction

Consumers	Reservation Price	Winning Bid
1	\$40	\$20.01
2	20	
3	15	
4	90	60.01
5	60	50.01
6	50	40.01
Total consumer surplus		69.96
Total seller surplus		170.04
Total available surplus		240

TABLE 13.2

Posted Price

Consumers	Reservation Price	Price Paid
1	\$40	\$40
2	20	
3	15	
4	90	40
5	60	40
6	50	40
Total consumer surplus		80
Total seller surplus		160
Total available surplus		240

tion, sellers get a lower percentage of this by using a posted price (\$160 versus \$170 + ε). Auction usage increases seller surplus because of its price discriminatory property. The mechanism itself gets consumers to reveal their reservation prices.

VALUE OF INFORMATION

The preference-revealing nature of auctions not only guarantees buyers with the highest reservation prices purchase the good or service; it also gets buyers to identify themselves when the demand is unknown. Managers can use auctions to gather more information about demand before announcing prices.

Repurchase Tender Offers

An example of this use of auctions is **repurchase tender offers (RTOs)**. An RTO is used by managers to buy back stock shares from current shareholders. Because shareholders are not required to sell, to induce selling, RTOs generally offer a premium above the current market price. Until 1981 just about every RTO was fixed in price. Managers would announce the buyback price per share and wait to see shareholder demand at that price. Since 1981 modified Dutch auction RTOs have been the mechanism of managerial choice. And it is easy to see why: Relative to a fixed-price RTO, the modified Dutch variety generally acquires shares at a lower total cost. Several studies show that the average premium paid for fixed-price tendered shares is 15 to 20%, whereas those in modified Dutch auctions average 10 to 15%.

Unlike a fixed-price RTO, where one price is announced, in a modified Dutch one, managers announce a price range at which they are willing to repurchase tendered shares. Generally the minimum price is set at a slight premium over market, while the maximum approaches that of a fixed-price RTO. Any shareholder willing to tender reveals his or her valuations as the seller sends managers a pricing schedule (how many shares to tender at \$ X). Managers use the schedules of individual shareholders to construct a market supply schedule. They then determine how many shares they need and set the share price. All sellers who value the shares below that price receive the stated price. Again, note the timing differential between the fixed-price and modified Dutch auction RTOs. In the fixed-price offer, managers set a price before knowing the supply schedule; in the modified Dutch auction, they set the price after seeing the supply schedule.

Citizens First Financial Corporation (CFFC) is a savings and loan holding company with headquarters in Bloomington, Illinois. The company has roughly \$325 million in assets. Its managers wanted to repurchase 391,000 shares of its stock from shareholders. On October 31, 2000, managers announced a modified Dutch auction RTO. (The share price that day was \$14.) They set a price range of \$15 to \$17 per share—a premium above the market of 7.1 to 21.4%. Shareholders had until December 1 to tender their shares. On December 11, 2000, CFFC

Repurchase tender offers (RTOs)

Offers used by managers to buy back stock shares from current shareholders.

managers announced they had purchased 391,096 shares at a price of \$16 (a premium of 14.3% above the market).

Let's look at the situation as CFFC managers may have viewed it. Many times managers repurchase shares because they believe the stock is undervalued. An RTO is used by managers to distribute company value to shareholders. RTOs annually distribute over \$4 billion to shareholders. Assume CFFC managers believe their stock will be worth \$20 per share in the near future. They want to buy back some shares now but do not know the valuations of shareholders. After stating their beliefs, managers construct Table 13.3 to summarize them. It shows how many shares will be tendered at different prices in three possible supply scenarios—strong, medium, and weak—and the generated value associated with each price–supply scenario. The profit equals the difference between a tender price and the expected future value of \$20.

If managers choose a fixed-price RTO, they must set the price before knowing the supply schedule. They might choose a price based on an expected value (*EV*)

$$EV(\$15) = \$2,000,000(0.40) + \$1,550,000(0.30) + \$1,400,000(0.30) = \$1,685,000$$

$$EV(\$16) = \$1,660,000(0.40) + \$1,600,000(0.30) + \$1,260,000(0.30) = \$1,522,000$$

$$EV(\$17) = \$1,800,000(0.40) + \$1,245,000(0.30) + \$1,200,000(0.30) = \$1,453,500$$

STRATEGY SESSION: The Use of Sophisticated Pricing within an Auction Format

Priceline.com is a good example of how managers can use sophisticated pricing policies within an auction format. The company was launched in April 1998 after Priceline managers helped develop software for a computer reservation system. The software allowed airline managers to make real-time adjustments to prices based on various factors, like unoccupied seats or competitor prices. Each airline sees only its own prices, but managers at Priceline can see prices across all member airlines.

Airline seats are perishable goods: Once the plane departs, empty seats are worthless because they cannot be used again. And the marginal cost of flying a plane with one more passenger is almost zero, so any price is profitable for the airline. As Brian Ek of Priceline.com states, "The airlines fly with up to 700,000 empty seats a day. Naturally they would love to sell those seats if they could without affecting their retail fare structure." Priceline.com developed an auc-

tion mechanism to allow airlines to do just that. The company has sold over 5 million airline tickets since 1998.

Priceline.com uses a reverse auction mechanism. In reverse auctions, buyers name the price they are willing to pay for a good or service. The seller then decides whether to accept or reject this price. Priceline's auction operates as follows: A consumer specifies a departure date, the departing and destination airports, the price the consumer is willing to pay for a ticket, and a credit card number. All sales are final. If Priceline finds a ticket at or below that price, the consumer is obligated to purchase it. After receiving a consumer's price, Priceline examines its database to determine if any airline is offering tickets at or below that price. If there are tickets, Priceline buys them. The profit to Priceline is the difference between what the consumer is willing to pay for the ticket and the price Priceline was charged.

TABLE 13.3

Shareholder Supply Schedule

Price	Strong # of Shares	Profit	Medium # of Shares	Profit	Weak # of Shares	Profit
\$15	400,000	\$2,000,000	310,000	\$1,550,000	280,000	\$1,400,000
16	415,000	1,660,000	400,000	1,600,000	315,000	1,260,000
17	600,000	1,800,000	415,000	1,245,000	400,000	1,200,000
Probability of shareholder's willingness to tender						
	0.40		0.30		0.30	

If CFFC managers had used such an analysis, they might have set a tender price of \$15.

How might managers have increased firm value by structuring the RTO as a modified Dutch auction? If they used the auction, they would not select a tender price without knowing the supply schedule. What was the expected value of the

Although Priceline claims it has increased the market power of consumers (because they are free to name their own prices), many disagree with this claim. Basically Priceline's reverse auction allows airlines to practice price discrimination. Like airlines, Priceline sells tickets on the same airplane at different prices depending on the prices quoted by consumers. Say I submit a price of \$300 for a ticket from New York to Chicago. If my friend submits a price of \$250, she would buy an identical product at a cheaper price (assuming airlines were willing to sell tickets for \$250). Priceline's reverse auction gets consumers to name their reservation prices, and Priceline then charges them these prices. So the reverse auction does not guarantee that consumers receive a product at the lowest price; it simply gives them the chance to purchase a product at their reservation prices.

Priceline's price-discriminating auction also improves on traditional price discrimination schemes

by making discounts less transparent to both consumers and rivals. Airlines need not post any special rates, which reduces the probability that rivals could engage in a disastrous price war. Also, the company can practice price discrimination selectively. For example, there is some evidence that Priceline will accept lower prices from first-time customers. When those consumers subsequently bid the same price for the same product, they find their bids are rejected.

Airlines that work with Priceline also receive demand information that helps them in their pricing decisions. The company compiles all bidding information (both successful and unsuccessful) and gives this information to airlines weekly. As Ek states, "They can see all of the demand, every consumer's price offer for every route, going all the way down to \$1. . . . It's a great way for the airlines to privately move more inventory that was likely to go unsold through their retail sales channels."

auction strategy? We use expected value because when CFFC managers announced the Dutch RTO on October 11, they did not know the supply schedule. So, if the willingness to tender was strong, they would announce a price of \$15; if it was medium, they would announce a price of \$16; and if it was weak, a price of \$15 would be set. Hence the expected value of a modified Dutch auction RTO would be

$$\begin{aligned} EV(\text{auction}) &= \$2,000,000(0.40) + \$1,600,000(0.30) + \$1,400,000(0.30) \\ &= \$1,700,000 \end{aligned}$$

Managers are generally better off using a modified Dutch auction RTO. They get shareholders to reveal their valuations and hence can buy back shares at a lower price than if they used a fixed price. This creates value for the remaining shareholders because some shares are retired at a lower cost. And the expected number of shares tendered is

$$0.4(400,000) + 0.3(400,000) + 0.3(280,000) = 364,000$$

RISK AVERSION

As we will see in Chapter 14, most individuals are risk-averse. What are the effects of risk aversion on bidding behavior? In second-price auctions, risk preferences do not influence bidding strategy. Bidders in these auctions should always bid up to their reservation prices. However, risk preference affects bidding behavior in first-price auctions.

Consider the choices facing a risk-averse bidder in a first-price auction. The prevailing uncertainty is this: Will my bid win the auction? Because first-price auctions have no dominant strategies, managers must anticipate the bids of others and hence can only partially control this uncertainty. Their controlling mechanism is the bid price itself. A higher bid increases the probability of winning; a risk-averse bidder will pay to avoid loss, so he or she simply raises the bid by some amount of “bidding insurance.” Thus risk-averse people bid higher than risk-neutral bidders. Their rush toward certainty is tempered, however, because higher bids also reduce the surplus.

If bidders are risk-averse, risk-neutral managers can increase the revenues generated through the use of auction rules. If bidders are likely to be risk-averse, then managers should use a first-price auction since this type of auction will elicit higher bids.

What if the roles were reversed—bidders are risk-neutral and sellers are risk-averse? How should this affect managerial actions? Here too the seller prefers first-price auctions. Although we know that the expected revenues from the four auction formats are equal, there is more risk in second-price auctions. That is, even though the revenue distribution means are equal, there is greater spread in possible revenues for second-price auctions. Because risk-averse managers prefer to avoid risk, they should use first-price auctions.

NUMBER OF BIDDERS

We have seen that in all auction formats, the expected bid is given by the reservation price of the second-highest bidder. Therefore managers should consider actions to increase these prices, just as they do in market settings. One possible variable that seems easily controlled is the number of bidders.

Consider how markets work. As the number of sellers increases (other factors held constant), the equilibrium price is pushed downward toward the marginal cost (or the seller's reservation price). And in perfectly competitive markets, the long-run equilibrium price is identical to the marginal cost. These markets have many sellers offering the same good or service to consumers. Auctions are the reverse of this: many buyers and one (or few) sellers. What works on the supply side also works on the demand side. The entry of more bidders (buyers) into a market must push the demand curve rightward and hence increase the price. So the more bidders a manager can induce to enter the auction, the greater is the expected surplus for the seller.

Are 1,000,001 bidders better than 1,000,000? No, that is not quite the case, although it may be true. Auctions are efficient under most conditions, so even a moderate number of bidders will make the seller's revenue approach the expected maximum valuation (conditioned on the probability function). Managers do not want to pay much to attract the 1,000,000th bidder to an auction, but the first 20 or so are important. The following example assumes the reservation prices of bidders are uniformly distributed, with the highest price being \$100. A seller's expected revenue in English auctions is then given by

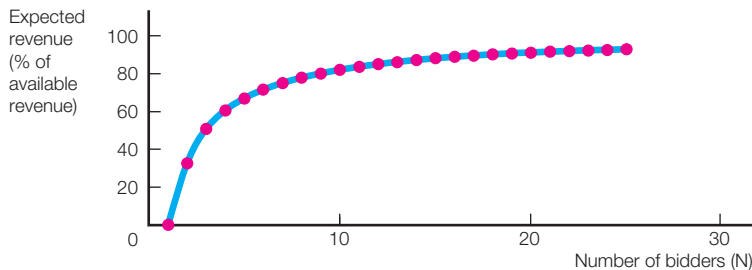
$$b = [(N - 1)/(N + 1)](\text{Mean reservation price of bidders})$$

where N = the number of bidders. Figure 13.2 shows how a seller's expected revenue increases as the number of bidders increases.

FIGURE 13.2

Expected Revenue versus Number of Bidders

As the number of bidders increases, the expected revenue of the seller increases.



Intuitively, the expected revenue of the seller increases because of the greater competition caused by additional bidders. In placing bids, managers must consider the trade-off between the probability of winning the auction and the level of surplus (reservation price — bid price) they will realize. A lower bid results in higher surplus but a decreased probability of winning: One loses an auction if only one person bids higher. As more bidders join the auction, the probability of someone bidding higher increases quickly. Hence as more bidders join the auction, bids must approach the reservation prices of bidders (and seller revenue increases). Note too that a seller's expected revenue approaches its maximal level with fewer than 30 bidders.

WINNER'S CURSE

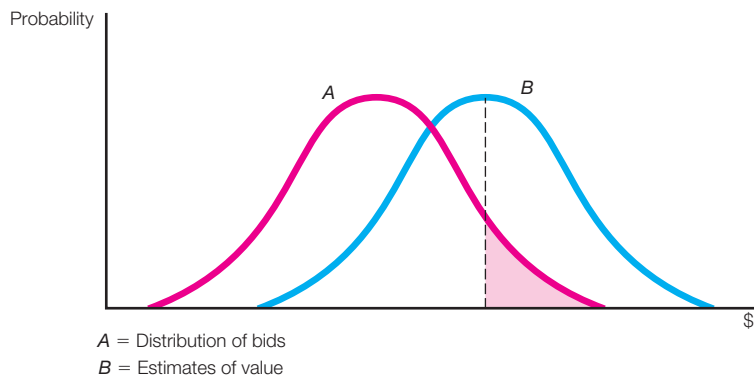
In some auctions the value of the object or service being auctioned off is not known with certainty, although it has a common value to all bidders. For example, the U.S. government auctions the mineral rights to tracts of federally owned land. When the seller uses a sealed-bid, first-price auction in these situations, bidders are exposed to what is known as the **winner's curse**. The winner's curse is a decision bias that has proven very costly to some managers. Since managers are unsure of an item's true value, they must estimate its value in formulating a bidding strategy. Those managers who estimate a high value are more likely to submit a high bid. So managers who overestimate the true value have a higher probability of winning the auction, but their bid may be higher than the value of the item. So, they win the auction, but they pay more than the item's worth. Hence, they fall prey to the winner's curse. Let us see why this is true.

Winner's curse When the bidder with the highest estimated value bids the most and wins the auction, but the bid amount may exceed the true value of the object.

FIGURE 13.3

The Winner's Curse

If the value of the item being auctioned is uncertain, bidders with extreme estimates of the value can bid higher than the item's true value.



All bidders face the same decision problem: They must estimate the value of the object without knowing the estimates of others. Suppose each bidder makes an estimate, and on average, the estimates are approximately correct. Then the distribution of values might resemble that shown in Figure 13.3. An unbiased estimate of the true value, if managers knew the bids of others, would be the mean shown by the dashed line. Given this distribution of values across bidders, the distribution of bids would lie to the left, say, as distribution A. As you can see, bidders with extreme estimates (in the right tail of A) bid values that exceed the best estimate of the true value. The key is that they do not know their estimates are extreme (because they do not know the estimates of others). Hence they are likely to win the auction but pay more than the object's true value.

The winner's curse is a robust phenomenon; it has been documented in many situations (see Strategy Session: The Winner's Curse in Bidding for Oil Rights).

STRATEGY SESSION: The Winner's Curse in Bidding for Oil Rights

During their careers, many managers find themselves in competitive bidding situations. These might entail bidding for the services of a qualified job applicant, for production inputs, to acquire another company, or to obtain a contract for corporate or personal services. When the bidding is conducted as a sealed-bid auction and it involves an object of common value, managers must be aware of the winner's curse.

An example of how costly this curse can be is shown in the accompanying table, which compiles the actual bids (in millions of dollars) of oil companies for the right to drill for oil on two tracts of government-owned land. Bids are listed in descending order of magnitude.

The table clearly shows the range of estimates of how much oil is present within these tracts of land. For tract 1, the winning bid is 84% higher than the second-highest bid; for tract 2, this percentage increases to 181%. If we examine the ratio of the high bid to the low bid in each auction, we see that for tract 1, the high bid is almost 10 times the low bid; for tract 2, this ratio is 109.

The natural tendency for managers is to increase their bids when more bidders enter the auction. But

in these types of auctions, managers need to reduce their bids, not increase them. Obviously the manager who submitted the winning bid for tract 2 did not follow our recommendations. In both auctions, the winning bidder suffered the winner's curse. Neither firm realized a positive return from these tracts of land.

Bids to Drill for Oil on Federal Land (in Millions of Dollars)

Tract 1	Tract 2
32.5	43.5
17.7	15.5
11.1	11.6
7.1	8.5
5.6	8.1
4.1	5.6
3.3	4.7
	2.8
	2.6
	0.7
	0.7
	0.4

STRATEGY SESSION: Why Managers Must Care About Profit, Not Market Share

Is there a winner's curse outside of common values auctions? Perhaps and it's called winning the market share race.

Coke and Pepsi have been battling for years in the carbonated soft drink market. Coke was once supreme. Pepsi overtook Coke. Coke regained number one. And now Diet Coke has become number two.

The significance on the good news side is two-fold. From Coke's perspective, it's another victory over Pepsi. From the obesity crisis perspective, it's a victory of no calories/no sugar over mega calories/sugar. Diet Mountain Dew and Diet Dr Pepper also rose in the rankings.

But there's bad news too. The size of the carbonated soft drink market is shrinking as consumers switch to zero calorie water, energy drinks, caffeinated

non-carbonated drinks, and other potable liquids. The carbonated market fell 0.5% last year and 2.1% the previous year. The decline over time has been such that the market is now the same size as it was in 1996.

If winning the market share race is profitable, it's a good thing. But being the valedictorian of a bad school may seem good in relative terms, but it may not put you in a great position in terms of absolute intelligence. Likewise, winning the market share race may be bittersweet if it is not profitable.

Source: Christina Cheddar Berk, "Why Diet Coke's Victory Is Bittersweet for Coca-Cola," March 17, 2011, at www.cnbc.com/id/42132085.

Although sophisticated algorithms for bidding behavior in these types of auctions are available, they are too complex to discuss in this text. However, managers should consider the following issues when contemplating their bids:

- What is your information relative to that of other bidders about the value of the asset? The less information managers have relative to others, the more they need to lower their bids.
- How confident are you in your estimate of the object's true value? The less confident managers are, the more they should lower their bids.
- What is the number of bidders? The more bidders in the auction, the more managers should lower their bids.

CONCERNS IN AUCTION DESIGN

In designing auctions, managers must realize that no auction design best suits all situations. Managers need to carefully consider the interface between auction design and resulting incentives. Two issues managers need to address in choosing an auction's design are the ability of bidders to collude within the auction and the attractiveness of the auction to potential bidders.³

3. See P. Klemperer, "What Really Matters in Auction Design," *Journal of Economic Perspectives*, 16(2002), pp. 169–189.

The ability of bidders to collude reduces the expected revenue of the seller and basically overpowers the efficiency of auctions. Bidders can collude in many ways. One common way is for a group of bidders to form a *ring* in which the bidders do not bid against each other. A member of the ring is designated as the bidder, and this one person bids on the objects being auctioned. After the auction, the members of the ring meet and distribute the objects from the auction. Many times the form of the distribution is an auction among the members themselves.

Collusion is also possible in multiunit simultaneous auctions. In these auctions, multiple units of a good are auctioned off simultaneously. In the early stages of the auction, participants who collude signal each other to show what goods they desire. The colluding participants then do not bid up the prices of the identified objects. Klemperer claims this occurred in the auctioning of spectrum rights in Germany. Two large telecoms—Mannesman and T-Mobile—split the 10 available blocks. Each company acquired exactly half of the available blocks for the identical low price.

Managers also need to make their auctions attractive to potential bidders. If an auction does not attract enough bidders, the revenue raised by it can be lower than predicted. For example, if it is clear to bidders that one bidder will win the auction, others may not enter because they are sure they will not win. This happened in Glaxo's purchase of Wellcome. Though other firms were interested in Wellcome, none entered the bidding process because it was clear to them (and reinforced by Glaxo statements) that Glaxo would top any bid.

Bidders also may be deterred from entering an auction if the seller sets a reserve price that is too high or low. Setting a price that is too low can actually encourage collusion among bidders. If the price is set too low and there is a strong bidder, that bidder may find it easier (and cheaper) to collude with others to keep the price low. Basically, if the bidder does not collude, he or she faces the prospect of outbidding all other bidders, which is usually more expensive than colluding.

SUMMARY

1. There are four types of auction mechanisms: English, Dutch, and first- and second-price sealed-bid auctions.
2. In private-value auctions with symmetric, risk-neutral bidders and independent signals, the auction mechanism chosen makes no difference. All mechanisms generate the same expected revenues.
3. There is no need for managers to consider the bidding strategies of others in English and second-price auctions. In English auctions, managers should bid up to their reservation prices. In second-price auctions, managers should bid their reservation prices.

4. Dutch and sealed-bid auctions have no dominant strategy. Optimal behavior in these auctions is conditioned on beliefs about the strategies of others. In each auction, managers must consider the reservation price of the second-highest bidder. Then managers need to bid just above this price.

5. When selling goods, managers should choose reserve prices that reflect this guideline:

Optimal reserve price = Value of the object being auctioned off, if it does
not sell + (Managerial estimates of the highest reservation price divided by 2)

6. Auctions are often more efficient than posted-price schemes because they can better discriminate among consumers with differing reservation prices. Most auction designs ensure that the consumer with the highest reservation price wins the auction.

7. Auctions are useful if demand is unknown. Managers can use them to induce consumers to reveal their preferences before managers set a price. By knowing the demand before setting a price, managers increase revenues.

8. When bidding in sealed-bid, common-value auctions, managers need to recognize the tendency to bid too high. This tendency results in the winner's curse, where the high bidder wins the auction but pays more than the object's true value.

wwnorton.com/studyspace 

PROBLEMS

1. Consultant.com is a company that employs business professors as virtual consultants who supply answers to other companies' problems. Consultant.com wants to raise funds with a private equity issue. Unfortunately, because of fluctuations in the stock market, it is uncertain about the demand for its offering. It hopes to issue the stock at either \$45 or \$50. The demand is categorized into four possible scenarios. The following table shows demand for each scenario–price combination along with the beliefs regarding the probability of each possible state. Consultant.com must pay 10% of the generated funds to the investment bank that helped it identify potential investors. The company wants to maximize the funds raised.

Price/Share	State 1	State 2	State 3	State 4
\$45	1,750	1,975	2,220	2,445
\$50	1,200	1,415	2,001	2,305
Probability	0.35	0.20	0.30	0.15

What is the expected value of the stock offering if Consultant.com sets its price without knowing the future demand state? If Consultant.com can determine the future demand state by using a modified Dutch auction, what is its expected profit? If someone approached Consultant.com and told managers she could predict the future demand state, how much would that information be worth to them?

2. Your company is planning to auction off a manufacturing plant in Asia. You are asked to determine the auction design that will generate the highest revenue for the company. You believe that bidders will value the plant independently. Which design would you choose, and why?
3. There are 100 bidders in an English auction. A random sample of 40 bidders shows the following reservation prices:

Number of Bidders	Reservation Price
1	\$10
3	\$20
6	\$30
5	\$40
8	\$50
6	\$60
7	\$70
3	\$80
1	\$100

Assume the bidding distribution is normal.

- a. What is the mean value of bids across the 100 bidders?
- b. What is the probability of a bid being less than \$80?
4. Your company is bidding for a service contract in a first-price, sealed-bid auction. You value the contract at \$12 million. You believe the distribution of bids will be uniform, with a high value of \$16 million and a low value of \$3 million. What is your optimal bidding strategy with
 - a. 5 bidders?
 - b. 10 bidders?
 - c. 20 bidders?
5. The Philadelphia Eagles of the National Football League build a stadium. One revenue source during the construction of the stadium is a personal seat license (PSL), a one-time, up-front payment charged to season ticket holders before the stadium is built. It gives the buyer the right to purchase tickets for a particular seat.

The Eagles are uncertain about demand for seats in the new stadium. They have selected three price points for PSLs (\$6,000, \$7,000, and \$8,000). Management also has estimated that demand for PSLs could be low, medium, or high. Their beliefs are reflected in this table:

Price	Low Demand, Probability = 0.4	Medium Demand, Probability = 0.35	High Demand, Probability = 0.25
\$6,000	24,500 PSLs sold	28,000 PSLs sold	40,000 PSLs sold
\$7,000	21,500 PSLs sold	24,000 PSLs sold	32,000 PSLs sold
\$8,000	17,500 PSLs sold	22,000 PSLs sold	25,000 PSLs sold

Some of the stadium funding is provided by the city of Philadelphia. Because Eagles fans view PSLs as an attempt to take away consumer surplus, they resent them and have put pressure on the city government to limit their use. Therefore, the city has set a target of 25,000 seats assigned to PSLs. If the Eagles sell fewer than 25,000 PSLs, the city will grant the team a tax benefit of \$10/seat for each seat under 25,000. If the Eagles sell 25,000 or more PSLs, no tax break will be given. A consulting group has told the Eagles that the team is better off using a modified Dutch auction to sell the PSLs. The group has estimated the cost of running the auction at an additional \$5.1 million. Eagles management has come to you for help. The managers want to know whether they should use an auction and what the expected benefits will be. What will you tell them?

6. Your company is bidding for a broadband spectrum license. You are asked to submit an optimal bidding strategy. You expect that bidders will have independent private values for the licenses because each bidder presently has a different structure in place. You believe the valuations for these licenses will be between \$200 million and \$700 million. Your own valuation is \$650 million. There is some uncertainty about the auction design that will be used, so you must suggest an optimal bidding strategy for the following auction designs:
 - a. Second-price, sealed-bid auction.
 - b. English auction.
 - c. Dutch auction.

Please describe your strategy. Is it a dominant strategy?

EXCEL EXERCISE: AUCTIONS

Suppose that a firm's shares are currently selling for \$65/share. The firm believes that the shares are undervalued and that their true value is \$80/share. They think it would be a good investment for the company to buy back its shares (of course, at a price less than \$80). But how to buy back the shares and at what price?

The firm is contemplating two methods. One is called a posted price. The firm announces that it will buy back its shares at price $\$X/\text{share}$. It sometimes sets a limit on the number of shares they will buy, but it almost always reserves the right to buy more. Shareholders decide whether they wish to sell their shares (all or part) at that price. The firm announces the buyback price before they ever see a share of tendered stock.

The second method is the modified Dutch auction. Here the shareholders announce to the firm not only the number of shares they are willing to tender, but also their reservation price. They are willing to accept any price higher than that. The firm only announces the price at which it will buy back the stock after it has observed what the market response to its offer to buy back shares has been.

Suppose in the case of posted price, the firm has estimated the following table of the number of shares it estimates will be tendered at each of the buyback prices it is considering. It is also not sure what the market response to its buy-back offer will be. It feels that it could be strong (a 30% chance), medium (a 50% chance), or weak (a 20% chance). It has estimated the number of shares (in thousands) that it anticipates will be tendered under each buyback price and demand response scenario (shown in the table below). The firm is risk-neutral and wishes to maximize its profit from buying back its stock.

Buy Back Price	Demand Response		
	Strong Probability = 0.3	Medium Probability = 0.5	Weak Probability = 0.2
78	18	16	15
76	16	15	7
74	14	13.8	4
72	12	10	3
70	10	6	2

Note that the share entries make qualitative sense in that, for any given buy-back price, the number of shares estimated tendered decreases as the demand response gets weaker, and that, for any given demand response, more shares are tendered as the buyback rises. For example, the number of shares that would be tendered at 72 under a strong demand include the 10(000) that would be tendered at price 70, that is, at the increased price of 72, the higher price has enticed 2(000) more shares to be tendered. The 14(000) shares estimated to be tendered at price 74 includes the 10(100) tendered at price 70 and the 2(000) tendered at price 72, and an additional 2(000) enticed when the price is raised to 74.

Let's see what you would do under the posted price case.

Call up your spreadsheet. Enter 80 in cell A1. Then click on the lower right-hand corner of cell A1 with your mouse and drag down until cell A5. Column A is the firm's reservation price of the stock.

Enter 78 in cell B1, 76 in cell B2, and so on, until you've added 70 in cell B5. Column B is the column of contemplated buyback prices.

Enter $=A1-B1$ in cell C1. Then click and drag until cell C5. Column C is your profit per share if you buyback the shares at the respective column B prices.

Enter 18 in cell D1, 16 in column D2, and so on, until you enter 10 in column D5. Column D is the number of estimated shares tendered at each price under the strong demand scenario.

Enter $=C1*D1$ in cell E1 and click and drag until cell E5. Column E is the expected profit of the firm at each buyback price under the strong demand scenario.

Enter 16 in cell F1, 15 in column F2, and so on, until you enter 6 in column F5. Column F is the number of estimated shares tendered at each price under the medium demand scenario.

Enter $=C1*F1$ in cell G1 and click and drag until cell G5. Column G is the expected profit of the firm at each buyback price under the medium demand scenario.

Enter 15 in cell H1, 7 in column H2, and so on, until you enter 2 in column H5. Column H is the number of estimated shares tendered at each price under the weak demand scenario.

Enter $=C1*H1$ in cell I1 and click and drag until cell I5. Column I is the expected profit of the firm at each buyback price under the weak demand scenario.

What is the expected profit of the firm under each buyback price?

Enter $=0.3*E1+0.5*G1+0.2*I1$ in cell J1 and click and drag until you reach cell J5. Column J is the column of expected profit if the firm chooses the corresponding buyback price. Search column J for the highest value or enter $=\text{Max}(J1:J5)$ in cell J6. The best result will be 73.6 in cell J4. Reading leftward in the fourth row, you will see that this entails choosing a buyback price of 72.

Now let's view the modified Dutch auction, in which the buyback price is set after, rather than before, the demand response is known.

Enter $=\text{Max}(E1:E5)$ in cell E6. This will give 100 in cell E6. Then enter $=\text{Max}(G1:G5)$ in cell G6. This will give 82.8 in cell G6. Then enter $=\text{Max}(I1:I5)$ in cell I6. This will yield 30 in cell I6. This tells us that if the demand response is strong, the firm will choose a buyback price of 70, since the profit of 100 is the highest expected under the strong demand response. The firm will choose a buyback price of 74 under the medium demand scenario, since the profit of 82.8 is the highest expected under the medium demand response. And the firm will choose a buyback price of 78 under the weak demand response, since the profit of 30 is the highest expected under the weak demand response.

So knowing the response results in different buyback prices depending on the demand response, as opposed to the same buyback price, regardless of the demand response in the posted price scenario. Note that, in this case, the optimal posted price (72) is never chosen under any demand scenario under the modified Dutch case.

Note too that none of these share numbers may be correct when the shareholders submit their shares and their reservation prices. But the decision as to which method to choose has to be made first, and that decision occurs based on the best available information at the time. Note too that it's possible for the two methods to result in the same buyback price (if the highest profit in any demand scenario has a price that is the same as the posted price).

Which one is better, that is, which yields the highest expected profit? The modified Dutch auction does, because it picks the highest expected profit in each column, as opposed to the posted price, which picks the highest expected profit for a set price.

What is the expected profit under the modified Dutch auction? Enter $=0.3*E6+0.5*G6+0.2*I6$ in cell A7. That equals 77.4. Since the best posted price gives expected profits of 73.6, the modified Dutch auction increases expected profits by 3.8 ($= 77.4 - 73.6$).

PART 7

RISK, UNCERTAINTY, AND INCENTIVES

CHAPTER 14

LEARNING OBJECTIVES

Risk and Probability
Probability Distributions and Expected Values
Comparisons of Expected Profit
Road Map to Decision
The Expected Value of Perfect Information
Measuring Attitudes toward Risk: The Utility Approach
Attitudes toward Risk: Three Types
The Standard Deviation and Coefficient of Variation: Measures of Risk
Adjusting the Valuation Model for Risk
Certainty Equivalence and the Market for Insurance
Summary
Problems
Excel Exercise: Expected Utility

RISK ANALYSIS

As you can see, the managerial world interacts with others. In that sense, it mirrors life. Managers encounter risk and uncertainty, and they face decisions characterized by two conditions: a strategic nature and incomplete information. We have discussed strategy in previous chapters.

In incomplete information settings, managers do not possess all the relevant information. Hence in having made a decision, a manager cannot be sure of its outcome. Chance, or the actions of others, may play a role. For example, marketing a new product involves risk because managers are uncertain of the level of demand. If demand is high, profit is greater; if demand does not justify break-even volume, profit is negative. Can managers be certain the investment will reap rewards? Usually not: When they make the launch decision, the future is still uncertain.

Managers face many decisions where the tantalizing possibility of high reward is offset by higher risk. Drilling a well for oil or investing in alternative energy is risky. In making decisions, managers must form beliefs about the future. The inability to see the future can paralyze managers. Managers need to use decision tools to see through the maze of possibilities and identify sensible decisions.

In this chapter we present a variety of tools to help managers improve their decision making. We start with the concept of expected value, which summarizes a set of possible outcomes into a single representative value. We then look at decision trees. Many decisions involve myriad possible outcomes, many of which are conditional on pure chance or the actions of others. For example, a manager

might wonder, “If interest rates are low and my competitor keeps prices high, then demand for my new product will be high—but what if interest rates rise, competitors’ prices fall, or new entrants come into the market?” Decision trees give us a visual and intuitive guide through the web of possible consequences and allow us to structure decisions in a simple, sequential way.

We also examine techniques to reduce uncertainty. Techniques and databases can improve our ability to forecast the future. For example, the designer of a new product can conduct market research on consumer acceptance and analyze the performance of comparable products to provide more information about the likely demand. An oil company can use geological models and satellite surveys to give a clearer picture of a drilling site and its likelihood of holding oil. In short, we can invest in information about the future. We show how valuable such information is and how its quality depends on its reliability.

The other major concept we introduce is that of expected utility. People react differently to risk, much as they have differing preferences for real goods. Given a choice between a safe investment with a low rate of return and a risky investment with a high expected return, some would choose the former, others the latter; this reflects their tolerance for risk. We see how expected utility analysis can reflect risk tolerance in managerial decisions.

RISK AND PROBABILITY

In ordinary parlance, **risk** is a hazard or a chance of loss. If managers of a biotechnology firm invest \$2 million in research and development knowing there is a substantial chance they will lose their money if a successful product is not launched, they are making a risky investment. Moreover, the bigger the chance of loss or the greater the size of a potential loss, the more risky a particular course of action is.

To analyze risk, it is necessary to define **probability**. Suppose a situation exists in which one of a number of possible outcomes can occur. For example, if a gambler throws a single die, the number that comes up may be 1, 2, 3, 4, 5, or 6. A probability is the number attached to each possible outcome. It is the proportion of times this outcome occurs over the long run if this situation is repeated. The probability that a particular die will show a 1 is the proportion of times this will occur if the die is thrown many, many times; the probability that the same die will show a 2 is the proportion of times this will occur if the die is thrown many, many times; and so on.

If a situation is repeated a large number of times, R , and if outcome A occurs r times, the probability of A is

$$P(A) = \frac{r}{R} \quad (14.1)$$

Thus if a die is “true” (meaning that each of its sides is equally likely to come up when the die is rolled), the probability of its coming up a 1 is $1/6$, or 0.167,

Risk Hazard or chance of loss.

Probability The likelihood or chance that something will happen.

STRATEGY SESSION: Pfizer and Its New Cholesterol Drug

Pharmaceutical company managers face decisions of great risk when developing new drugs. We illustrate how managers can use expected value by considering the following issues confronted by the managers of drug company Pfizer.

Pfizer managers had invested \$1 billion developing a new, potential blockbuster drug that could add tens of billions of dollars to the company's market value. As with all drugs, managers knew there was a risk of unwanted consequences. Side effects from any drug can potentially cause injury or death, which can result in costly lawsuits. Indeed, several drug companies have lost a large part of their market value recently in such suits (consider Merck's drug Vioxx). Managers use clinical trials to help them assess the risk of new drugs. In 2006 Pfizer managers were performing trials in which some 15,003 people at high risk of cardiovascular disease were either taking the new drug or were allocated to a control group taking the existing drug Lipitor.

The trials showed that 82 people taking the new drug died compared with 51 people taking the control drug. Because of the high existing cardiovascular risk of the subjects, managers had expected some mortality. However, in evaluating the effectiveness and safety of the new drug, they needed to know whether the additional deaths were purely a statistical anomaly or whether they indicated that the new drug was indeed more risky.

The protocol managers established in advance for the trials was that if the probability that additional deaths could be attributed to the new drug exceeded 1 in 100, the trials would be stopped. A statistical analysis of the trial results showed that this threshold had been passed. In other words, managers could be 99% sure the results were not anomalous and did indeed reveal the new drug to be more deadly than the control drug.

The decision facing Pfizer managers involved a trade-off. If they stopped the trials, they would lose an opportunity to earn high profits. If they continued the trials and sought regulatory approval to produce and sell the new drug, the firm might face costly lawsuits from people who attributed their injury or death to an allegedly defective drug. How could Pfizer managers map out and quantify the range of outcomes arising from each decision (abandon the drug or not)? Moreover, how could Pfizer managers measure the likelihood of reaching one of these future outcomes? And after managers discussed the possible outcomes and established their beliefs, how could they organize the data to reach a sensible decision that was in the best interests of the firm and its shareholders?

In reality, Pfizer managers decided to abandon the trials, and they lost their \$1 billion investment. As a result of this decision, the firm's market value fell by \$21 billion.

because if it is rolled many, many times, this will occur in one-sixth of the rolls of the die.

What we have just provided is the so-called **frequency definition of probability**. However, in some situations this concept of probability may be difficult to apply because the situations cannot be repeated over and over. In 2008 Kentucky Fried Chicken was testing a new grilled chicken product to decide whether to bring it to market in early 2009.¹ This is an experiment that cannot be repeated under essentially the same circumstances. Market and other conditions vary from

Frequency definition of probability An event's limit of frequency in a large number of trials.

1. Yahoo! News, March 24, 2008.

STRATEGY SESSION: What are the major global risks in 2011?

The World Economic Forum (well known for its yearly economic summit in Davos) each year issues a listing and analysis of what it considers to be the most pressing global risk. These risks will shape the environment in which business will operate.

*“Two risks are especially significant given their high degrees of impact and interconnectedness. **Economic disparity** and **global governance failures** both influence the evolution of many other global risks and inhibit our capacity to respond effectively to them.”*

The ranking of players in the world economy is shifting but, despite this rebalancing, there is still evidence that economic disparity is increasing. Moreover, there is also clear disparity between countries on how sustainable growth can be achieved. In an increasing number of countries, the response takes the form of nationalist and populist movements.

The Forum goes on to identify three important clusters of risks

- *“**The “macroeconomic imbalances” nexus:** A cluster of economic risks including macroeconomic imbalances and currency volatility, fiscal crises and asset price collapse arise from the tension between the increasing wealth and influence of emerging economies and high levels of debt in advanced economies. Savings and trade imbalances within and between countries are increasingly unsustainable while unfunded liabilities create extreme long-term pressure on fiscal positions. One way out of these imbalances would be coordinated global action but this is challenging given the conflicting interests of different states.”*
- *“**The “illegal economy” nexus:** This nexus examines a cluster of risks including state fragility, illicit trade, organized crime and corruption. In a networked world, governance failures and economic disparity create opportunities for such illegal activities to flourish. In 2009, the value of illicit trade around the globe*

was estimated at US \$1.3 trillion and growing. These risks, while creating huge costs for legitimate economic activities, also weaken states, threatening development opportunities, undermining the rule of law and keeping countries trapped in cycles of poverty and instability. International cooperation—both on the supply side and on the demand side—is urgently needed.”

- *“**The “water-food-energy” nexus:** A rapidly rising global population and growing prosperity are putting unsustainable pressures on resources. Demand for water, food, and energy is expected to rise by 30–50% in the next two decades, while economic disparities incentivize short-term responses in production and consumption that undermine long-term sustainability. Shortages could cause social and political instability, geopolitical conflict, and irreparable environmental damage. Any strategy that focuses on one part of the water-food-energy nexus without considering its interconnections risks serious unintended consequences.”*

Finally, the report identifies four more risk classes to watch out for:

- *“**Demographic challenges** adding to fiscal pressures in advanced economies and creating severe risks to social stability in emerging economies.”*
- *“**Resource security** issues causing extreme volatility and sustained increases over the long run in energy and commodity prices, if supply is no longer able to keep up with demand.”*
- *“**Retrenchment from globalization** through populist responses to economic disparities, if emerging economies do not take up a leadership role.”*
- *“**Weapons of mass destruction**, especially the possibility of renewed nuclear proliferation between states.”*

Source: World Economic Forum, “Global Risks 2011,” Sixth Ed., eds Kristel Van der Elst and Nicholas Davis. {Quotations are from the Executive Summary.}

month to month. If KFC were to delay the introduction, the types of fast-food meals available from other firms (like McDonald's or Burger King), the prices of these competing products, the advertising campaigns of other firms, and a host of other relevant factors would probably change.

In dealing with situations of this sort, managerial economists sometimes use a **subjective definition of probability**. According to this definition, the probability of an event is the degree of confidence or belief on the part of the manager that the event will occur. If the manager believes that outcome *X* is more likely to occur than outcome *Y*, the probability of *X* occurring is higher than the probability of *Y*. If the manager believes that it is equally likely that a particular outcome will or will not occur, the probability attached to the occurrence of this outcome equals 0.50. We use this concept of probability to model managerial beliefs.

Subjective definition of probability The degree of a manager's confidence or belief that the event will occur.

PROBABILITY DISTRIBUTIONS AND EXPECTED VALUES

If all possible outcomes are listed and the probability of occurrence is assigned to each outcome, the resulting table is called a **probability distribution**. For example, suppose managers at Adept Technology, a San Jose, California, manufacturer of robots, think the probability is 0.6 that they can develop a new type of robot in one year and the probability is 0.4 they cannot do so in this length of time. The probability distribution is as follows:

Event	Probability of Occurrence
New robot is developed in one year	0.6
New robot is not developed in one year	<u>0.4</u>
	1.0

Note that the probabilities sum to 1, which must be the case if all possible outcomes or events are listed.

If Adept Technology will earn a profit of \$1 million if managers develop the new robot in one year and lose \$600,000 if they do not, we can readily calculate the probability distribution of profit from the new robot:

Profit	Probability
\$1,000,000	0.6
-\$ 600,000	0.4

Moreover, we can also calculate the expected value of the profit

$$\$1,000,000(0.6) + (-\$600,000)(0.4) = \$360,000$$

Probability distribution A table that lists all possible outcomes and assigns the probability of occurrence to each outcome.

The expected value is the weighted average of the profit corresponding to the various outcomes, each of these profit figures being weighted by its probability of occurrence.

In general, expected profit is expressed by the equation

$$\text{Expected profit} = E(\pi) = \sum_{i=1}^N \pi_i P_i \quad (14.2)$$

where π_i is the level of profit associated with the i th outcome, P_i is the probability that the i th outcome will take place, and N is the number of possible outcomes. Because $N = 2$, $\pi_1 = \$1,000,000$, $\pi_2 = -\$600,000$, $P_1 = 0.6$, and $P_2 = 0.4$ in the case of Adept Technology, equation (14.2) says precisely the same thing as the equation that precedes it.

COMPARISONS OF EXPECTED PROFIT

To decide which of a number of courses of action to take, managers can compare the expected profits. For example, suppose managers of the Jones Corporation, a producer of automobile tires, are thinking of raising the price of their product by \$1 per tire. The managers estimate the following. If they raise the price, they will realize a profit of \$800,000 if the current advertising campaign is successful, and they will incur a \$600,000 loss if the campaign is not successful. Managers believe there is a 0.5 probability the current advertising campaign will be successful and a 0.5 probability it will not succeed.

Under these circumstances, the expected profit to the firm if managers increase price equals

$$\$800,000(0.5) + (-\$600,000)(0.5) = \$100,000$$

As indicated, the expected profit is the sum of the amount of money gained (or lost) if each outcome occurs times the probability of occurrence of the outcome. In this case there are two possible outcomes: (1) The firm's current advertising campaign is successful or (2) it is unsuccessful. If we multiply the amount of money gained (or lost) if the first outcome occurs by its probability of occurrence, the result is $\$800,000(0.5)$. If we multiply the amount of money gained (or lost) if the second outcome occurs by its probability of occurrence, the result is $-\$600,000(0.5)$. Summing these results, we get \$100,000, which is the expected profit if the firm raises its price.

What would be the expected profit if managers at the Jones Corporation did not increase price? Suppose they believe that with no price increase, profit will reach \$200,000. For simplicity, we assume that this profit level is regarded as certain if the price is not increased. If managers want to maximize the expected profit, they should not increase price because the expected profit equals \$200,000 if price is not increased but only \$100,000 if it is. Later in this chapter, we discuss at length

the circumstances under which it is rational to maximize the expected profit—and how to proceed if it is not rational to do so.

ROAD MAP TO DECISION

Managers frequently face choices over alternative strategies and the payoff of the chosen strategy will depend on the actions of others. Because it is not easy to anticipate precisely the behavior of others, managers will form beliefs about the future. To analyze any such scenario, a decision tree is useful.

Decision tree A diagram that helps managers visualize their strategic future.

A **decision tree** is a diagram that helps managers visualize their strategic future. It represents the situation as a series of choices, each of which is depicted by a fork (sometimes called a juncture or branching point). A decision fork represents a choice by which managers must commit to a strategy. A chance fork represents a point at which chance influences the outcome. To differentiate between a decision fork and a chance fork, we place a small square at the former juncture but not at the latter.

In Figure 14.1 we show the decision tree for the situation facing managers at the Jones Corporation. Managers are considering a price increase. Starting at the left side of the diagram, Jones managers must choose whether to increase price. We represent this choice as a decision node, so it is square. If managers do not institute a price increase, the consequence is certain: Managers earn a profit of \$200,000. Therefore, \$200,000 is shown at the end of this branch. If the managers do increase price, one of two consequences will occur. If the current advertising campaign is successful, customers accept the price increase, and managers earn \$800,000. This is shown in the upper branch. The lower branch represents the future where the campaign is not successful, customers do not accept the price increase, and managers lose \$600,000. On the basis of history, 50% of similar campaigns were successful, and 50% were not.

The tree illustrates the various options and gives managers an easy way to compare payoffs across strategies. We can then determine more easily what strategies managers should take. Begin at the right side of the decision tree, where the profit figures are located. The first step is to calculate the expected profit when managers are situated at the chance fork immediately to the left of these payoff figures. In other words, this is the expected profit to managers given that chance will influence which subsequent branch is followed. Because there is a 0.50 probability of following the branch culminating in a profit of \$800,000, and a 0.50 probability of following the branch culminating in a loss of \$600,000, the expected profit at this chance fork is

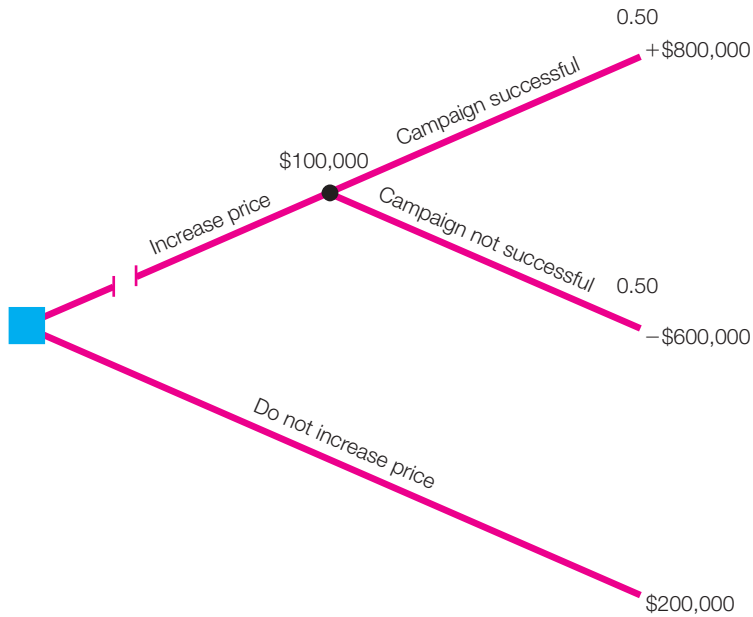
$$0.50(\$800,000) + 0.50(-\$600,000) = \$100,000$$

This number is written above the chance fork in question to show that this is the expected profit when the firm is located at that fork. Moving farther to the left

FIGURE 14.1

Decision Tree, Jones Corporation

If the Jones Corporation increases its price, the expected profit is \$100,000. If it does not increase its price, the expected profit is \$200,000.



along the decision tree, it is clear that managers have a choice of two branches; one leads to an expected profit of \$100,000, and the other leads to a \$200,000 expected profit. If managers want to maximize their expected profit, they should choose the latter branch. In other words, they should not increase price. Because the former branch (increase price) is nonoptimal, we place two vertical lines through it.

Of course this graphic procedure for analyzing the Jones Corporation's pricing problem amounts to precisely the same thing as the calculations we made in the previous section. In short, the decision tree lets you visualize complex decision problems. In the Jones case, there were only two courses of action and only two chance possibilities. However, when a decision problem gets more complicated with several choices and each choice has several possible outcomes, it is easy to get lost in the intricate maze of possibilities. A decision tree provides a simple visual map that brings order and transparency to such complex decision structures.

PROBLEM SOLVED: Should Managers at Genco Exploration Drill?

One important business in which the concepts just presented have been applied is natural gas and oil production from shale rock formations. Very large amounts of money are being invested in this sector. Managers use these analytical tools as an aid to decision making. To illustrate how these concepts are applied, consider the case faced by managers of Genco Exploration who had to decide whether to drill a well at a site in Pennsylvania. The managers had information concerning the cost of drilling and the price of gas and oil as well as geologists' reports concerning the likelihood of striking significant deposits. The geologists' reports led the managers to believe that if a well were drilled, there was a 0.60 probability of finding no deposits, a 0.15 probability of finding the equivalent of 10,000 barrels, a 0.15 probability of finding 20,000 barrels, and a 0.10 probability of finding 30,000 barrels.

Although these probabilities are helpful, they do not, by themselves, help managers choose whether to drill. Managers need to consider the payoff attached to each scenario. Suppose managers estimate they will realize a \$90,000 loss if they find no oil, a \$100,000 profit if they find 10,000 barrels of oil, a \$300,000 profit if they find 20,000 barrels, and a \$500,000 profit if they find 30,000 barrels. Based on these beliefs, should the managers drill the well?

Managers can answer this question by constructing the decision tree shown in the following figure. Starting at the left side of the diagram, the first choice is up to the managers who can either

drill the well or not. If the branch representing not drilling is followed, the expected profit is zero, which is shown at the end of this branch. (Why? Because the firm neither gains nor loses if it does not drill.) If the branch representing drilling the well is followed, we come to a chance fork: It is uncertain whether the well will strike oil and, if it does, how much oil it will produce. The highest branch following this chance fork represents the consequence that no oil is found, in which case managers lose \$90,000, shown at the end of this branch. The next branch following this chance fork represents the consequence that 10,000 barrels are found, in which case the firm gains \$100,000, shown at the end of this branch. Similarly, the lower branches following this chance fork represent the consequences that 20,000 and 30,000 barrels are found; the number at the end of each of these branches is the corresponding profit to the firm.

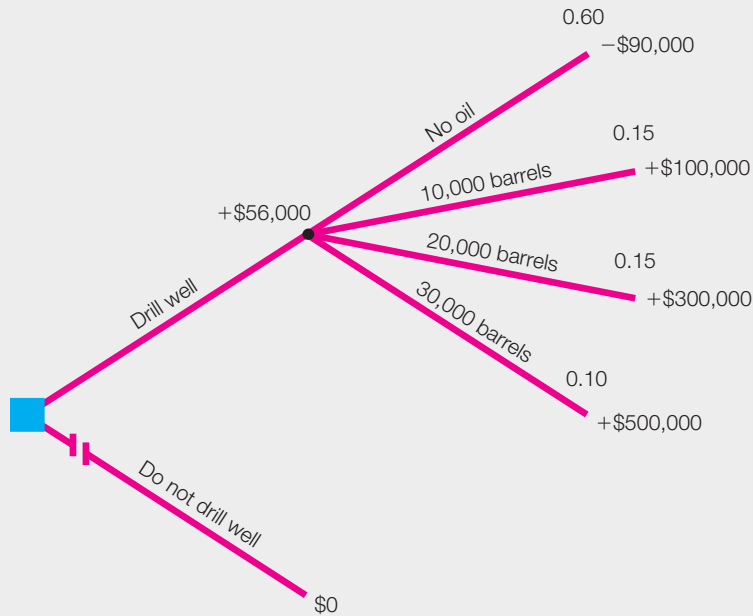
Having constructed this decision tree, the firm's managers can compute the expected profit to the firm if it is situated at the chance fork immediately to the left of the profit (or loss) figures. If the firm is at this fork, there is a 0.60 probability that the branch culminating in a \$90,000 loss will be followed, a 0.15 probability that the branch culminating in a \$100,000 profit will be followed, a 0.15 probability that the branch culminating in a \$300,000 profit will be followed, and a 0.10 probability that the branch culminating in a \$500,000 profit will be followed. To calculate the expected profit if the firm is situated at this fork, the firm's managers should multiply each possible value of profit (or loss)

THE EXPECTED VALUE OF PERFECT INFORMATION

Sometimes managers can obtain information that will dispel uncertainty (to varying extents). If managers can buy such information, how much should they be willing to pay for it? We define the value of perfect information as the increase in expected profit if the manager can obtain completely accurate information con-

Decision Tree, Genco Exploration

If Genco Exploration drills the well, the expected profit is \$56,000. If it does not, the expected profit is zero.



by its probability, and sum the results. The expected profit if the firm is situated at this chance fork equals

$$0.60(-\$90,000) + 0.15(\$100,000) + 0.15(\$300,000) + 0.10(\$500,000) = \$56,000$$

In the figure, this result is written above the chance fork in question to show that this is the expected profit if the firm is located at that fork.

Going farther along the decision tree to the left, managers have a choice of two branches; one leads to an expected profit of \$56,000, and the other leads to an expected profit of zero. So if managers want to maximize expected profit, they should choose the former branch—that is, drill the well.

cerning future outcomes. Managers generally pay for this information prior to knowing it. For example, managers at the Jones Corporation would find it worthwhile to accurately estimate how successful their advertising campaign will be.

To illustrate how managers can compute the expected value of perfect information, we return to the Jones case. We begin by evaluating the expected monetary value to managers of access to perfectly accurate information of this

sort. If managers can obtain perfect information, choosing the optimal decision becomes easy. If the campaign is going to be successful, managers will increase price and earn a profit of \$800,000. If it is not going to be successful, managers will not increase price and will earn \$200,000. To managers with access to perfect information, the expected profit is

$$0.50(\$800,000) + 0.50(\$200,000) = \$500,000$$

Even though the managers can obtain perfect information, it is not revealed to them until after they have paid for it. We can imagine asking some forecaster to give her expert (and never wrong) opinion as to whether the advertising campaign will succeed. There is a 0.50 probability that this forecaster will tell them that the advertising campaign will be successful, in which case managers can increase price and earn a profit of \$800,000. There is also a 0.50 probability the forecaster will tell managers the campaign will fail, in which case they would not increase price and would earn a profit of \$200,000. So, as shown, the expected profit if managers obtain access to a perfect forecast (which is revealed after payment) is \$500,000.

Recall that managers expect a profit of \$200,000 if they make a decision without perfect information, not \$500,000. The difference between these two amounts (\$500,000 – \$200,000, or \$300,000) is the expected value of perfect information. It shows the amount by which the expected profit increases as a consequence of managers having access to perfect information; \$300,000 is thus the managers' reservation price for this information.

In many circumstances it is important that managers think about the worth of perfect information. Managers are frequently offered information by testing services, research organizations, news bureaus, credit rating agencies, and a variety of other organizations. Unless managers know how much particular types of information are worth, it is difficult to decide rationally whether they should be bought. The value of information analysis presented in this section is useful to guide such decisions because it shows the maximum amount managers should be willing to spend to obtain perfect information. The calculation of what to pay for less-than-perfect information is more complex and is not presented here. Needless to say, this amount is less than what a manager would pay if the information were perfect. And when the accuracy of the information falls below a certain level, it is worthless to managers.

MEASURING ATTITUDES TOWARD RISK: THE UTILITY APPROACH

In discussing the behavior of managers at both Jones and Genco Exploration, we assumed that managers want to maximize expected profit. We now examine how risk affects managerial behavior. Imagine that a small business is offered the following choice:

PROBLEM SOLVED: Evaluating an Investment in a New Chemical Plant

To illustrate the usefulness of the expected value of perfect information, consider an actual case in which managers constructed a decision tree to determine whether a major U.S. corporation should invest in a new plant. The main product of the new plant was a brightener; but with new processing methods, a valuable by-product was produced as well. The exact amounts of both products produced were uncertain. Minute quantities of impurities in the raw materials used in the process could greatly influence the amounts of brightener and by-product produced. Also, there were uncertainties concerning the costs of raw materials and plant efficiency.

The table shows the expected value of perfect information concerning by-product quantity, impurities, raw material costs, and plant efficiency. As you can see, the critical uncertainties were those regarding by-product quantity and the level of impurities. For example, perfect information concerning by-product quantity was worth up to \$6.2 million. On the other hand, information regarding raw materials

and plant efficiency was much less important for this decision. Indeed, the expected value of perfect information concerning plant efficiency was close to zero. Based on these results, the managers decided to research the uncertainties regarding by-product quantity and the level of impurities before committing to the construction of the new plant.

**Expected Value of Perfect Information
Concerning Factors Influencing Whether to
Build New Chemical Plant**

Factor	Expected Value of Perfect Information (Millions of Dollars)
By-product quantity	6.2
Level of impurities	3.9
Raw material costs	0.3
Plant efficiency	0.0

1. A certain profit of \$2,000,000.
2. A gamble with a 50–50 chance of a \$4,100,000 profit or a \$60,000 loss.

The expected profit for the gamble is

$$0.50(\$4,100,000) + 0.50(-\$60,000) = \$2,020,000$$

so managers should choose the gamble over the certainty of \$2,000,000 if they want to maximize expected profit. However, it seems likely that many managers, especially those of small businesses, would prefer the certainty of \$2,000,000 because the gamble entails a 50% chance that the firm will lose \$60,000—a substantial sum for a very small firm. Moreover, many managers may feel they can do almost as much with \$2,000,000 as with \$4,100,000, and therefore the extra profit is not worth the risk of losing \$60,000.

Whether the firm's managers will want to maximize the expected profit in this situation depends on their attitude toward risk. If the amount at stake in a decision

STRATEGY SESSION: FedEx's and Ata Holdings Corporation's Disclosures about Risk

As is typical of all publicly traded companies after Sarbanes-Oxley, FedEx tells investors and potential investors about its exposure to market risk. Such disclosure was caused by the failure to provide information about the true exposure to risk in the Enron, Tyco, and WorldCom cases.

FedEx states, "We have no significant exposure to changing interest rates on our long-term debt because interest rates are fixed on the majority of our long-term debt." With respect to a \$500 million debt with a floating interest rate that was to mature in August 2007, FedEx stated that it did not employ interest rate hedging to mitigate the risks with respect to this borrowing. The company's reasoning was that a 10% increase in the interest rate on its outstanding floating-rate debt would not have a material effect on its results of operations. On May 31, 2007, it had an estimated fair value \$2.4 billion in outstanding fixed-rate, long-term debt. The market risk for such debt (estimated to be \$36 million as of May 31, 2006) was estimated as the potential decrease in fair value resulting from a 10% increase in interest rates.

FedEx also disclosed its risk due to currency fluctuations. Most of its transactions were denominated in U.S. dollars, but it had significant transactions in the euro, Chinese yuan, Canadian dollar, British pound, and Japanese yen. FedEx stated that "distribution of our foreign currency denominated transactions is that such currency declines in some areas of the world are often offset by currency gains in other areas of the world." In fact, during 2006 and 2007, FedEx believed that operating income was positively impacted due to foreign currency fluctuations. However, favorable foreign currency fluctua-

tions also may have had an offsetting impact on the price it obtained or the demand for its services. On May 31, 2007, a 10% strengthening of the value of the dollar relative to the above-mentioned currencies would decrease operating income by \$151 million in 2008. Note the performance of sensitivity analysis (the hypothetical 10% changes in interest rates and exchange rates).

Finally, FedEx disclosed its risk associated with changes in fuel prices. This was mitigated by its use of fuel surcharges passing the increased fuel costs to customers. FedEx concluded that a 10% increase in the price of fuel would not materially affect its earnings. However, there is a lag in the deployment of its fuel surcharges, and should a material price change occur in fuel, FedEx disclosed that its operating income might be affected.

In addition, it discussed "many risks and uncertainties":

1. Preserving the strong reputation and value of the brand.
2. Threat of technology/Internet failure on operations.
3. Assurance of adequate physical capital to handle the business volume.
4. Threat of competitors.
5. Need to effectively operate, integrate, leverage, and grow acquired businesses.
6. Ability to acquire fuel at a reasonable price.
7. Need for FedEx Ground's personnel to be classified as independent contractors.
8. Role of increased security costs.
9. Role of the international regulatory environment.
10. Role of terrorist activities on operations/economic conditions.

11. Impact of new tax, accounting, labor, and environmental rules/laws.
12. Ability to manage its cost structure.
13. Ability to maintain good relations with employees.
14. Assuring a supply of high-quality labor.
15. Managing health care and other employee benefits.
16. Managing changes in volumes, customer demand parameters, and prices charged.
17. Innovating in service and growth of product lines.
18. Controlling legal fees and avoiding liability.
19. Adapting and responding to changes in technology that impact the demand for services.
20. Responding to adverse weather conditions/natural disasters that impact system operations.
21. Responding to a pandemic.
22. Assuring accessibility to financial capital at a reasonable price.

Consider now a company in poor financial shape in Chapter 11 bankruptcy (ATA holdings—an airline that terminated operations on April 3, 2008). In its 2004 10-K report, ATA spelled out the risks that might render its forward-looking statements about the expected future of the company incorrect:

- The ability to develop and execute a revised business plan for profitable operations, including restructuring flight schedules, maintaining the support of employees, and regauging the fleet of aircraft.
- The ability to develop, prosecute, confirm, and consummate a plan of reorganization with respect to Chapter 11 cases.

- Risks associated with third parties seeking and obtaining bankruptcy court approval to terminate or shorten the exclusivity period to propose and confirm one or more plans of reorganization, for appointment of a Chapter 11 trustee, or to convert one or more assets to a Chapter 7 case.
- The ability to obtain and maintain normal terms with vendors and service providers.
- The ability to maintain contracts that were critical to its operations.
- The potential adverse effects of the Chapter 11 reorganization on liquidity or results of operations.
- The ability to attract and retain customers.
- Demand for transportation in markets in which the company operated.
- Economic conditions.
- The effects of any hostilities or act of war.
- Salary costs.
- Aviation fuel costs.
- Competitive pressures on pricing (particularly from low-cost competitors).
- Weather conditions.
- Government legislation and regulation.
- Other risks and uncertainties listed from time to time in reports the company periodically filed with the Securities and Exchange Commission.

In fact the inability to maintain a military contract and skyrocketing fuel costs ultimately did ATA in.

Sources: FedEx, Form 10-K, Annual Report, Filed July 13, 2007, and ATA Holdings Corp. 10-K for December 31, 2004.

is large relative to the overall value of the firm and the managers' bonuses are likely impacted, they may be overwhelmed at the thought of taking a 50% chance of losing \$60,000, which could cost them their bonuses. On the other hand, if they manage a big corporation, the prospect of a \$60,000 loss may not be unsettling; and they may prefer the gamble to the certainty of a mere \$2,000,000 gain.

Fortunately we need not assume that managers want to maximize the expected profit. Instead we can construct a **utility function** for a manager that measures his or her attitude toward risk. (This concept of utility should not be confused with that discussed in Chapter 3. As we will see, it is a different concept.) From this utility function, we can identify the optimal strategy for managers conditional on their attitude toward risk.

Utility function Function used to identify the optimal strategy for managers conditional on their attitude toward risk.

Constructing a Utility Function

We define a rational manager as one who maximizes expected utility. That is, she does not want to harm herself. The manager chooses the strategy with the highest expected utility. But what (in this context) is utility? It is a value that is attached to all possible outcomes of the decision by the manager. The manager's utility function represents the level of satisfaction (or benefit or welfare) she attaches to each possible outcome. It is risk-adjusted, so it also represents risk preferences. How do managers compute **expected utility**? It is the sum of the utility of each outcome times the probability of the outcome's occurrence. For example, if a situation has two possible outcomes, *A* and *B*, if the utility of outcome *A* is 2 and the utility of outcome *B* is 8, and if the probability of each outcome is 0.50, the expected utility equals

Expected utility The sum of the utility of each outcome times the probability of the outcome's occurrence.

$$0.50(2) + 0.50(8) = 5$$

What is the expected utility if managers at Genco Exploration drill the well under the circumstances described on page 540? It equals

$$0.60U(-90) + 0.15U(100) + 0.15U(300) + 0.10U(500)$$

where $U(-90)$ is the utility that managers attach to a monetary loss of \$90,000, $U(100)$ is the utility attached to a gain of \$100,000, $U(300)$ is the utility attached to a gain of \$300,000, and so on. Because there is a 0.60 probability of a \$90,000 loss, a 0.15 probability of a \$100,000 gain, a 0.15 probability of a \$300,000 gain, and a 0.10 probability of a \$500,000 gain, the preceding formula shows the expected utility. What is the expected utility if managers do not drill the well? It equals $U(0)$ because under these circumstances, it is certain that the gain is zero.

To find the utility the manager attaches to each possible outcome, begin by ranking the utility attached to two levels of profit arbitrarily. The utility of the best consequence is set higher than that of the worst one. Managers in the oil-drilling problem might set $U(-90)$ equal to 0 and $U(500)$ equal to 50. Utility scales are discretionary. We can choose end points arbitrarily, but once we have chosen the scale,

we must stick with it. Therefore, managers could set $U(-90)$ equal to 1 and $U(500)$ equal to 10. It will make no difference to the ultimate outcome of the analysis.²

Managers should then consider payoffs that fall between these extremes. They need to consider choices between the certainty of one of the other possible levels of profit and a gamble where the possible outcomes are the two profit levels whose utilities we set arbitrarily. In the oil-drilling case, suppose managers want to find $U(100)$. To do so, managers should consider whether they prefer the certainty of a \$100,000 gain to a gamble where there is a probability P of a gain of \$500,000 and a probability $(1 - P)$ of a loss of \$90,000. They should consider at what value of P they are indifferent between the certainty of a \$100,000 gain and this gamble. Suppose this value of P is 0.40.

If the managers are indifferent between a certain gain of \$100,000 and this gamble, the expected utility of the certain gain of \$100,000 must equal the

2. The utility function we construct is not unique. Because we set the two utilities arbitrarily, the results vary, depending on the values of the utilities chosen. If X_1, X_2, \dots, X_n are the utilities attached to n possible monetary values, $(\alpha + \beta X_1), (\alpha + \beta X_2), \dots, (\alpha + \beta X_n)$ can also be utilities attached to them (where α and β are constants, and $\beta > 0$).

PROBLEM SOLVED: To Drill or Not

Once a manager estimates her utility function, it can help her decide whether to accept or reject particular gambles. Consider the actual case of Thomas Blair, president of Genco Exploration. Using the previous procedures, an economist estimated Blair's utility function. Suppose the result is as shown in Figure 14.2 and Blair must decide whether to drill the well described on page 540. He should drill the well if his expected utility if the well is drilled exceeds his expected utility if it is not drilled. As pointed out in the previous section, his expected utility if the well is drilled is

$$0.60U(-90) + 0.15U(100) + 0.15U(300) + 0.10U(500)$$

He can use this to estimate his utility as follows. Because $U(-90)$ equals 0, $U(100)$ equals 20, $U(300)$ equals 40, and $U(500)$ equals 50, his expected utility if the well is drilled is

$$0.60[0] + 0.15[20] + 0.15[40] + 0.10[50] = 14$$

It is important to note that utility functions measure utility of "wealth" or "net worth" and not the utility of "changes in income." For example, suppose a

person has wealth of \$100 and is facing this choice about gambling: If the coin comes up heads, he wins \$10, and he loses \$10 if it is tails. Managers can pose this choice in the following form

$$\begin{aligned} \text{Expected utility of gamble} &= \\ 0.5U(\$100 + \$10) + 0.5U(\$100 - \$10) \\ \text{versus} \\ \text{Expected utility of no gamble} &= 1.0U(\$100) \end{aligned}$$

With this in mind, managers at Genco should think of the values -90 , 100 , 300 , or 500 as the wealth or net worth of the firm under the different scenarios. In other words, these figures refer to the present value of future profit, and not simply the profit recorded in a single year.

If the well is not drilled, Blair's expected utility equals $U(0)$, which is 10 according to Figure 14.2. Therefore he should drill the well. Why? If he does not drill it, his expected utility is 10, whereas if he drills it, his expected utility is 14. Because he should maximize the expected utility, he should choose the action with the higher expected utility, which is to drill.

expected utility of the gamble. (Why? Because managers maximize expected utility.) Therefore,

$$U(100) = 0.40U(500) + 0.60U(-90)$$

And because we set $U(500)$ equal to 50 and $U(-90)$ equal to 0, it follows that

$$U(100) = 0.40(50) + 0.60(0) = 20$$

That is, the utility attached to a gain of \$100,000 is 20.

Managers now have the utility measures for three wealth levels; $U(-90)$ and $U(500)$ were set arbitrarily at 0 and 50 to determine the scale of the utility function. Then we derived that on this scale, $U(100)$ must equal 20.

Using the same procedure, managers of the oil company can estimate $U(300)$ and $U(0)$, the other utilities required to calculate the expected utility if they drill the well and the expected utility if they do not drill it. For example, to obtain $U(300)$, managers should consider whether they prefer a certain \$300,000 gain to a gamble where there is a probability P of a gain of \$500,000 and a probability $(1 - P)$ of a loss of \$90,000. Then managers should consider at what value of P they are indifferent between the certain \$300,000 gain and this gamble. Suppose this value of P is 0.80. Then the expected utility of a certain \$300,000 gain must equal the expected utility of this gamble

$$U(300) = 0.80U(500) + 0.20U(-90)$$

And because $U(500)$ equals 50 and $U(-90)$ equals 0, it follows that $U(300)$ equals 40.

Utility functions of managers represent the relationship between their utility and the amount of their profit (or loss). By evaluating $U(-90)$ (or $U(0)$), $U(100)$, $U(300)$, and $U(500)$ as in the previous paragraphs, managers can identify five points on their utility function, as shown in Figure 14.2. By repeated use of the procedure just described, managers can estimate many such points. (According to Figure 14.2, $U(0) = 10$.)

Risk averters When managers prefer a choice with a more certain outcome to one with a less certain outcome, when confronted with gambles offering equal expected wealth.

3. Consider a gamble in which there is a probability of P that the gain is π_1 and a probability of $(1 - P)$ that the loss is π_2 . A person is a risk averter if the utility of the gamble's expected profit, $U[P\pi_1 + (1 - P)\pi_2]$, is greater than the expected utility of the gamble, $PU(\pi_1) + (1 - P)U(\pi_2)$.

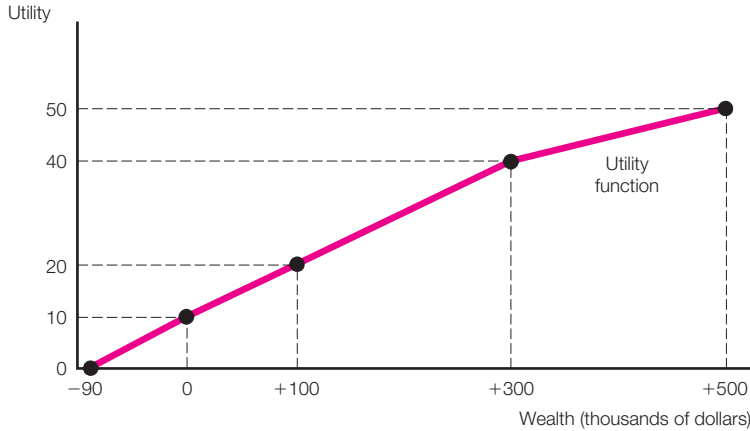
ATTITUDES TOWARD RISK: THREE TYPES

Although managers can expect that utility increases with monetary gain, the shape of the utility function can vary greatly depending on the preferences of the individual. Figure 14.3 shows three general types of utility functions. The one in panel A is like that in Figure 14.2 in the sense that utility increases with wealth but at a decreasing rate. In other words, an increase in monetary gain of \$1 is associated with smaller and smaller increases in utility as the wealth grows. Managers with utility functions of this sort are **risk averters**. That is, when confronted with gambles offering equal expected wealth, they prefer a choice with a more certain outcome to one with a less certain outcome.³

FIGURE 14.2

Utility Function

The decision maker's utility function is useful in indicating whether particular gambles should be accepted.



Panel B of Figure 14.3 shows a case in which utility increases with wealth at an increasing rate. In other words, an increase in monetary gain of \$1 is associated with larger and larger increases in utility as the wealth grows. Managers with utility functions of this sort are **risk lovers**. That is, when confronted with gambles offering equal expected wealth, they prefer a gamble with a less certain outcome to one with a more certain outcome.⁴

Finally, panel C shows a case where utility increases with wealth at a constant rate. In other words, an increase of \$1 in monetary gain is associated with a constant increase in utility as the wealth grows. Stated differently, utility in this case is a linear function of wealth

$$U = a + b\pi \quad (14.3)$$

where U is utility, π is wealth, and a and b are constants (of course $b > 0$). People with utility functions of this sort are **risk-neutral**.⁵ In other words, they maximize expected wealth, regardless of risk. It is easy to prove that this is true. If equation (14.3) holds,

$$E(U) = a + bE(\pi) \quad (14.4)$$

where $E(U)$ is expected utility and $E(\pi)$ is expected wealth.⁶ Consequently, because expected utility is directly related to expected wealth, it can be a maximum only when expected wealth is a maximum.

Risk lovers When managers prefer a gamble with a less certain outcome to one with a more certain outcome, when confronted with gambles offering equal expected wealth.

Risk-neutral When a manager maximizes expected wealth, regardless of risk.

4. A person is a risk lover if the utility of the gamble's expected profit, $U[P\pi_1 + (1 - P)\pi_2]$, is less than the expected utility of the gamble, $PU(\pi_1) + (1 - P)U(\pi_2)$.

5. A person can be a risk averter under some circumstances, a risk lover under different circumstances, and risk-neutral under still other circumstances. The utility functions in Figure 14.3 are "pure" cases in which the person is always only one of these types, at least in the range covered by the graphs.

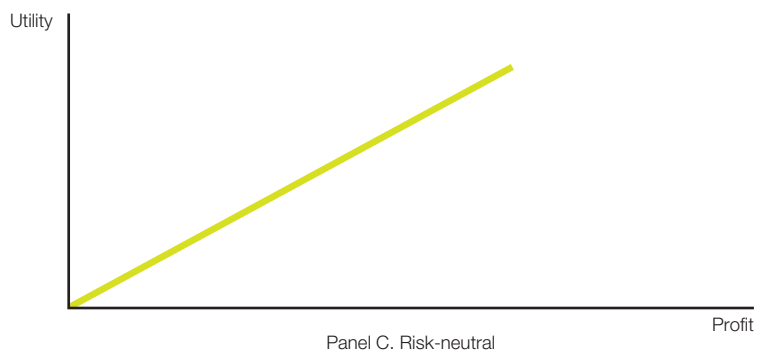
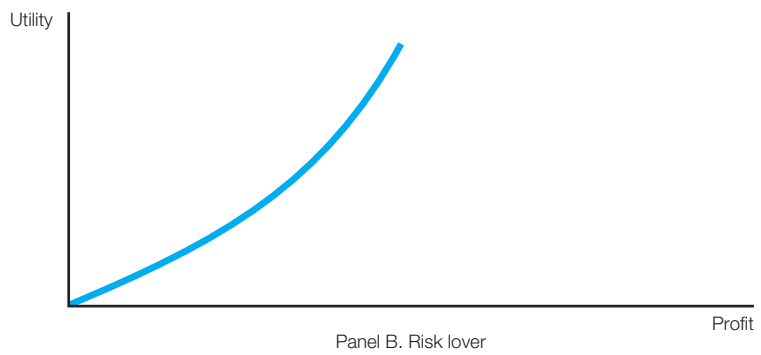
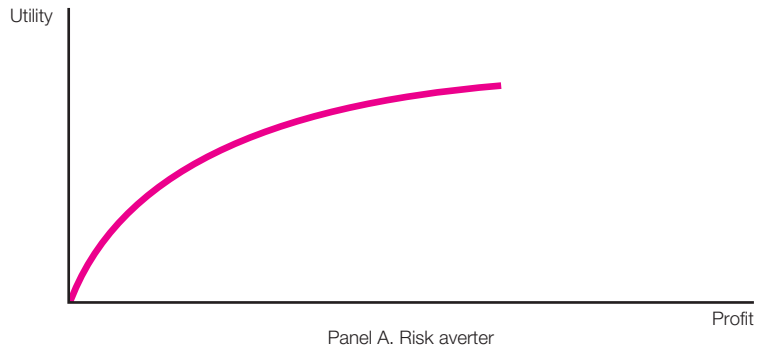
6. To illustrate that equation (14.4) is correct, suppose π can assume two possible values, π_1 and π_2 , and the probability that π_1 occurs is P and the probability that π_2 occurs is $(1 - P)$. Then, if $U = a + b\pi$,

$$\begin{aligned} E(U) &= P(a + b\pi_1) \\ &\quad + (1 - P)(a + b\pi_2) \\ &= a + b[P\pi_1 + (1 - P)\pi_2] \\ &= a + bE(\pi) \end{aligned}$$

because $E(\pi)$ equals $P\pi_1 + (1 - P)\pi_2$.

FIGURE 14.3**Three Types of Utility Functions**

Utility functions assume a variety of shapes. In panel A the decision maker is a risk averter; in panel B he or she is a risk lover; and in panel C he or she is risk-neutral.



THE STANDARD DEVIATION AND COEFFICIENT OF VARIATION: MEASURES OF RISK

Risk has many meanings. Some managers behave as if risk is measured in magnitude of negative outcomes. However managers can measure risk as being related to the dispersion of the probability distribution of profit resulting from the decision. For example, suppose Jones Corporation managers must decide whether to invest in a new plant. If the probability distribution of profit resulting from the new plant is as shown in panel A of Figure 14.4, the decision to invest in the new plant is more risky than the probability distribution shown in panel B. Why? Because the profit resulting from the new plant is more uncertain and variable in panel A than in panel B.

As a measure of risk, managers can often rely on the **standard deviation**, σ , the most frequently used metric for dispersion in a probability distribution.⁷ Managers need to consider all feasible choices and attach a payoff to each possible future. They should then input these to software that calculates the standard deviation. The software estimates the standard deviation of profit by computing the expected value of profit, $E(\pi)$ (recall equation (14.2)). It then subtracts this expected value from each possible profit level to obtain a set of deviations about this expected value. (The i th such deviation is $\pi_i - E(\pi)$.) Then we square each deviation, multiply the squared deviation by its probability of occurrence (P_i), and sum these products

$$\sigma^2 = \sum_{i=1}^N P_i [\pi_i - E(\pi)]^2$$

Taking the square root of this result, we obtain the standard deviation

$$\sigma = \left(\sum_{i=1}^N P_i [\pi_i - E(\pi)]^2 \right)^{0.5} \quad (14.5)$$

As an illustration, consider managers of a company who must decide whether to invest in new process technology. According to the company's engineers, there is a 0.3 probability that such an investment will result in a \$1 million profit, a 0.4 probability it will result in a \$0.2 million profit, and a 0.3 probability it will result in a \$0.6 million loss. Therefore the expected value of the profit from this investment is

$$E(\pi) = 0.3(1) + 0.4(0.2) + 0.3(-0.6) = 0.2$$

or \$0.2 million. Based on equation (14.5), the standard deviation is

$$\begin{aligned} \sigma &= [(0.3)(1 - 0.2)^2 + (0.4)(0.2 - 0.2)^2 + (0.3)(-0.6 - 0.2)^2]^{0.5} \\ &= (0.384)^{0.5} = 0.62 \end{aligned}$$

or \$0.62 million.

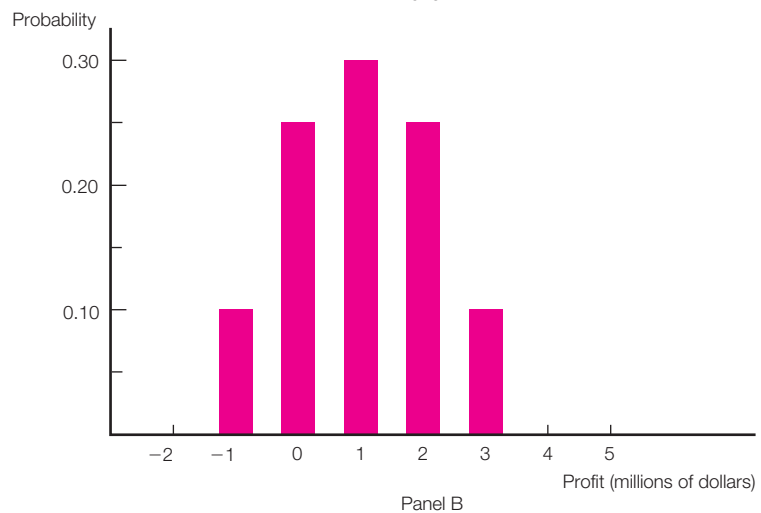
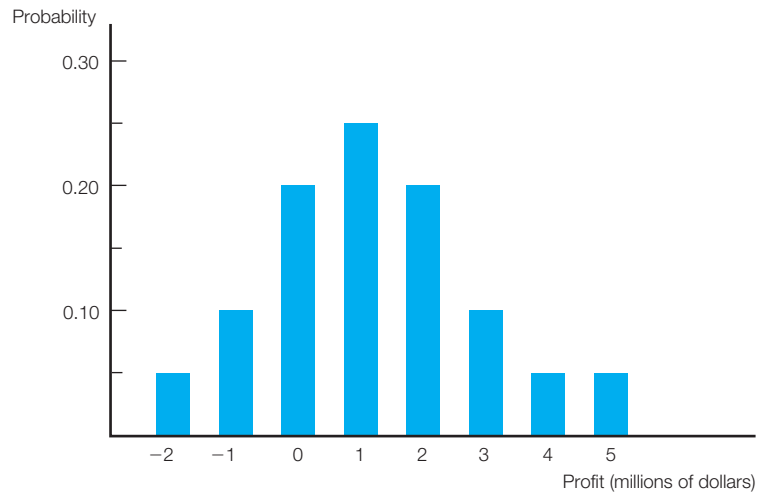
Standard deviation A measure of variation or dispersion of a payoff about its expected value; the most frequently used metric for dispersion in a probability distribution.

7. Although the standard deviation is often a useful measure of risk, it may not always be the best measure. Our discussion here and in subsequent sections of this chapter is necessarily simplified. The measures and techniques we describe are rough, but many analysts have found them useful.

FIGURE 14.4

Probability Distribution of the Profit from an Investment in a New Plant

The probability distribution in panel A shows more dispersion than that in panel B.



A larger standard deviation implies greater risk. If the standard deviation of the levels of profit resulting from the technology investment were \$2 million rather than \$0.62 million, there would be less certainty concerning its profitability. In other words, there would be more likelihood that its profitability would depart greatly from its expected value.

When managers use the standard deviation as a measure of risk, they implicitly assume the scale of the project is held constant. If one investment is twice as big as another, they should expect the standard deviations of the investments' profits to differ. Larger investments usually have greater standard profit deviations. To take account of the scale of the project, a measure of relative risk is required. Such a measure is the **coefficient of variation**, defined as

$$V = \frac{\sigma}{E(\pi)} \quad (14.6)$$

For example, in the case of the new technology investment, the coefficient of variation for the profit levels is $0.62/0.2$, or 3.1 .

Coefficient of variation

$$V = \frac{\sigma}{E(\pi)}$$

ADJUSTING THE VALUATION MODEL FOR RISK

According to the basic valuation model discussed in Chapter 1, managers must continually be concerned with the effects of their decisions on the present value of the firm's future profit, defined as

$$PV = \sum_{t=1}^N \frac{\pi_t}{(1 + i)^t} \quad (14.7)$$

But the firm's managers do not know with certainty what the firm's profit in future year t (that is, π_t) will be. The best they can do is use the expected profit (that is, $E(\pi_t)$) instead. How can they use the formula in equation (14.7) to account for risk?

One way is to use the **certainty equivalent approach**, which is related to the utility theory developed in the previous sections. For example, consider the manager of the firm considering an investment in process technology. Suppose she is indifferent between the certainty of a \$100,000 net worth (the net worth is the wealth of the firm) and the gamble involved in investing in the technology. We can now compare the certainty equivalent (\$100,000), with the expected profit (\$200,000). If the certainty equivalent is less than the expected net worth, the decision maker is a risk averter; if it is more than the expected net worth, she is a risk lover; and if it equals the expected net worth, she is risk-neutral.

Certainty equivalent approach

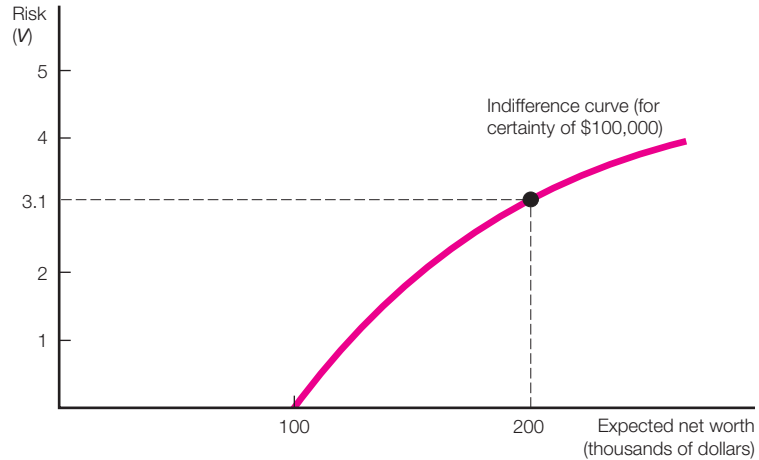
When a manager is indifferent about certainty and a gamble, the certainty equivalent (rather than the expected profit) can identify if the manager is a risk averter, risk lover, or risk-neutral.

On the basis of the decision maker's utility function, we can construct indifference curves of the sort shown in Figure 14.3. Each such indifference curve shows the certainty equivalent corresponding to various uncertain outcomes. Figure 14.5 shows that the manager is indifferent between the certainty of \$100,000 and a gamble in which the expected net worth is \$200,000 and the risk, as measured by the coefficient of variation, is 3.1 . Using such indifference curves, we can estimate the certainty equivalent of any uncertain situation. (In contrast to most of the indifference curves in Chapter 3, these indifference curves slope upward to the right. Why? Because the manager prefers less risk to more risk. In Chapter 3 the consumer preferred more of each commodity to less.)

FIGURE 14.5

Manager's Indifference Curve between Expected Profit and Risk

The manager is indifferent between gambles with the expected profit and risk shown here. Therefore, she is indifferent between the certainty of \$100,000 and a gamble in which the expected profit is \$200,000 and the coefficient of variation is 3.1. Similar indifference curves exist for riskless amounts other than \$100,000.



In practice, of course, it is not easy to obtain such indifference curves, just as it is not easy to obtain the utility functions on which they are based. Because managers have limited time and patience, it may not be feasible to get more than a limited amount of information concerning their utility functions. Nor is it always clear which of a number of managers is the relevant one. If many managers play an important role in a particular decision and they have diverse utility curves, they may come to different conclusions. But of course this should be expected. Indeed it would be strange if managers with various attitudes toward risk did not come to divergent conclusions when faced with a choice among alternatives entailing different amounts of risk.

CERTAINTY EQUIVALENCE AND THE MARKET FOR INSURANCE

Consider managers holding mortgage debt currently worth \$900 million in early 2009. They estimate there is a 25% chance the market will worsen and the worth of the bonds will drop to \$400 million. They also estimate a 75% chance the market will remain relatively constant. The expected value of the bonds is therefore

$$0.25(400) + 0.75(900) = 100 + 675 = \$775 \text{ million}$$

If the managers are risk-averse, they may have a utility (U) function of

$$U = W^{0.5}$$

Hence the expected utility from the bonds is

$$0.25(400)^{0.5} + 0.75(900)^{0.5} = 0.25(20) + 0.75(30) = 5 + 22.5 = 27.5$$

The certainty equivalent is the monetary sum that would make the managers indifferent between having that monetary sum for certain, and holding the bonds (which have an expected utility of 27.5)

$$U = W^{0.5} = 27.5$$

Squaring both sides yields $W = \$756.25$ million. This should be the reservation price of the managers for selling the bonds.

STRATEGY SESSION: PepsiCo Risk Management

PepsiCo faces market risks due to changes in commodity prices (raw materials and energy), foreign exchange, and interest rates. But it's not just the commodity price risk that differentiates PepsiCo from the risks facing FedEx. It's how it handles those risks.

PepsiCo uses derivatives as cash flow or fair value hedges. It limits hedging transactions to its underlying exposure. Therefore, a change in its derivative instruments would be substantially offset by an opposite change in the value of the underlying hedged items, thus protecting the company from risk due to those hedged items. Should the value of a hedge not offset the change in the value of the hedged item, the company would not have fully protected itself against risk. This is called an *ineffective hedge*.

In addition to the use of derivatives, PepsiCo protects against commodity price changes by contracts that fix prices and other pricing agreements and by diversifying geographically where it purchases commodities. In addition, it engages in productivity-enhancing endeavors with the objective of lowering its costs of doing business. The derivatives tend to be limited to no more than two years' duration and are used primarily in the natural gas, diesel fuel, and

fruit markets. PepsiCo has not suffered from ineffective hedges with either these instruments or the foreign currency instruments below.

Because 44% of net revenue is generated outside the United States (with slightly less than half of that from Canada, Mexico, and Great Britain), the company is subject to currency fluctuation risks. PepsiCo tends to enter into forward contracts of no more than two years' duration to handle this risk.

PepsiCo manages its debt and investment portfolios centrally, considering investment opportunities, tax consequences, and overall financing strategies. It uses interest rate swaps and cross-currency interest rate swaps to manage interest expense and foreign exchange risk. Such instruments change the interest rate and currency of specific debt issues. When a specific debt instrument is issued, a counterpart swap instrument is entered into concurrently. The notional amounts, interest payments, and maturity dates of the swaps match the principal amounts, interest rates, and maturity dates of the original debt instruments.

Source: 2007 PepsiCo Annual Report, pp. 38–39.

STRATEGY SESSION: The Use of Risk-Adjusted Discount Rates

Another way to introduce risk into the valuation model in equation (14.7) is to adjust the discount rate, i . This method, like that discussed in the previous section, is based on a manager's risk preferences. For example, suppose the figure shows a manager's indifference curve between expected rate of return and risk. As is evident from the fact that this curve slopes upward to the right, this manager is willing to accept greater risks only if he obtains a higher expected rate of return. Specifically, he is indifferent between a riskless investment yielding an 8% return and a risky investment ($\sigma = 2$) yielding an expected 12% return. In other words, as the risk increases, a larger expected profit is required to compensate for the higher risk.

The difference between the expected rate of return on a particular risky investment and that on a riskless investment is called the risk premium. For example, if the manager in the figure can obtain an 8% return from a riskless investment, he will require a risk premium of 4% (12% minus 8%) to compen-

sate for the level of risk corresponding to $\sigma = 2$. This is the extra rate of return required to induce him to make such a risky investment. If he is offered less than this 4% risk premium, he will not make the risky investment.

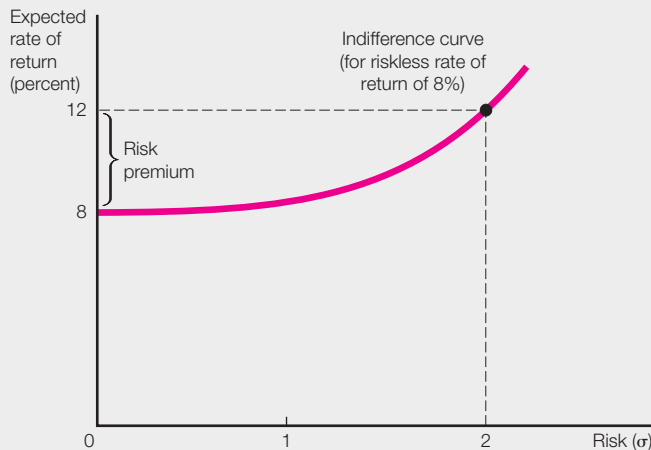
Because the required rate of return depends on how risky an investment is, managers can adjust the basic valuation model in equation (14.7) to account for risk by modifying the discount rate, i . The adjusted version of equation (14.7) is

$$PV = \sum_{t=1}^N \frac{\pi_t}{(1+r)^t} \quad (14.8)$$

where r is the risk-adjusted discount rate. The risk-adjusted discount rate is the sum of the riskless rate of return and the risk premium required to compensate for the investment's level of risk. If the risk is such that $\sigma = 2$, the risk-adjusted discount rate would be 12% for the manager in the figure. This risk-adjusted rate equals 8% (the riskless rate) plus 4% (the risk premium).

Manager's Indifference Curve between Expected Rate of Return and Risk

The manager is indifferent between a riskless return of 8% and gambles with the expected rate of return and risk shown in the figure. Similar indifference curves exist for riskless rates of return other than 8%.



Certainty equivalents are also used to create metrics in the insurance industry. The LBI Insurance Company provides coverage to protect individuals, such as the managers just described, against catastrophic events. LBI offers full coverage insurance—it will cover the insured's entire loss (\$500 in this case). LBI is risk-neutral. What premium should managers at LBI charge for this full coverage policy?

The expected payout for LBI managers is \$125. They have a 25% chance of paying out \$500 and a 75% chance of paying out \$0. Therefore, we expect a premium of at least \$125, so that LBI can expect to at least break even on the policy.⁸ What is the maximum that the mortgage-holding managers would pay for such a policy? This is where the certainty equivalent becomes relevant. If LBI charges \$143.75 for the policy, the managers would be left with \$756.25 ($\$900 - \143.75) for sure. (If no loss occurs, they pay only the premium, so they have $\$900 - \$143.75 = \$756.25$; but if the catastrophic event of 500 occurs, it is paid in full, in which case the managers would have $\$900 - \$143.75 - \$500 + \$500 = \$756.25$.) With the policy, the managers always will have the certain wealth of \$756.25 or the utility of $(\$756.25)^{0.5} = 27.5$. Notice that this is exactly the same expected utility they had with no insurance as shown earlier. Hence they are indifferent between buying the full coverage insurance policy for \$143.75 and facing the gamble on the bond market described earlier because both give an expected utility of 27.5.

The difference between the premium that will just cover LBI's expected payout (\$125) and the maximum that the managers would pay (\$143.75) is called the risk premium, or the amount the managers are willing to pay the insurance company above the expected value of the loss. The risk premium is \$18.75 in this case. If the premium (P) is between 125 and 143.75, the managers prefer to buy insurance (because the utility of $900 - P > 27.5$) and the insurance company will sell the policy (because the expected profit of $P - 125 > 0$). If LBI acts like a monopolist, it charges \$143.75 (or epsilon below) for the policy and extracts all the managers' surplus from purchasing it.

SUMMARY

1. The probability of an event is the proportion of times this event occurs over the long run. Expected profit is the sum of the amount of money gained (or lost) if each outcome occurs times the occurrence probability of the outcome.
2. Managers can use a decision tree to represent a decision problem as a series of choices, each of which is depicted by a decision fork. However, the tree also has chance forks reflecting the impact of risk on each choice. A decision tree can be used to determine the course of action with the highest expected profit. A variety of examples were discussed, including the decision by managers at Genco Exploration regarding whether they should drill a well in Pennsylvania.

8. We are simplifying here. Although the expected payout on this policy is \$125, LBI incurs other costs that it must cover if it is to stay in business, such as agents who sell policies and are paid commissions, back office staff who underwrite and process and pay claims, and executive compensation. In addition, we assume that LBI is risk-neutral. This is not an unreasonable assumption because LBI insures many individuals, and the probability of a catastrophic event occurring to its customers is close to 25% (if all are clones of the managers analyzed here). Thus LBI diffuses its risk by insuring many people.

3. The value of perfect information to managers is the increase in expected profit if they could obtain completely accurate information concerning the outcome (but they do not yet know what this information will be). This is the maximum amount managers should pay to obtain such information. Methods were provided to calculate the expected value of perfect information.

4. Risk is often measured by the standard deviation or coefficient of variation of the probability distribution of profit. How managers choose to maximize expected profit depends on their attitude toward risk. We represent a manager's attitude toward risk by his or her utility function.

5. To construct such a utility function, managers initially set the utility attached to two monetary values arbitrarily. Then managers choose between the certainty of one of the other monetary values and a gamble in which the possible outcomes are the two monetary values whose utilities were set arbitrarily. Repeating this procedure over and over can estimate a manager's utility function.

6. One way managers can adjust the basic valuation model for risk is to use certainty equivalents in place of the expected profit figures in equation (14.7). To do this, construct indifference curves (based on the manager's utility function) showing the certainty equivalent corresponding to various uncertain outcomes.

7. Another way managers can introduce risk into the valuation model is to adjust the discount rate. To do this, managers should construct indifference curves between expected rate of return and risk, based on their utility functions. Using such indifference curves, managers can estimate the risk premium (if any) that is appropriate.

PROBLEMS

1. The president of the Martin Company is considering two alternative investments, X and Y. If each investment is carried out, there are four possible outcomes. The present value of net profit and probability of each outcome follow:

Investment X			Investment Y		
Outcome	Net Present Value	Probability	Outcome	Net Present Value	Probability
1	\$20 million	0.2	A	\$12 million	0.1
2	8 million	0.3	B	9 million	0.3
3	10 million	0.4	C	6 million	0.1
4	3 million	0.1	D	11 million	0.5

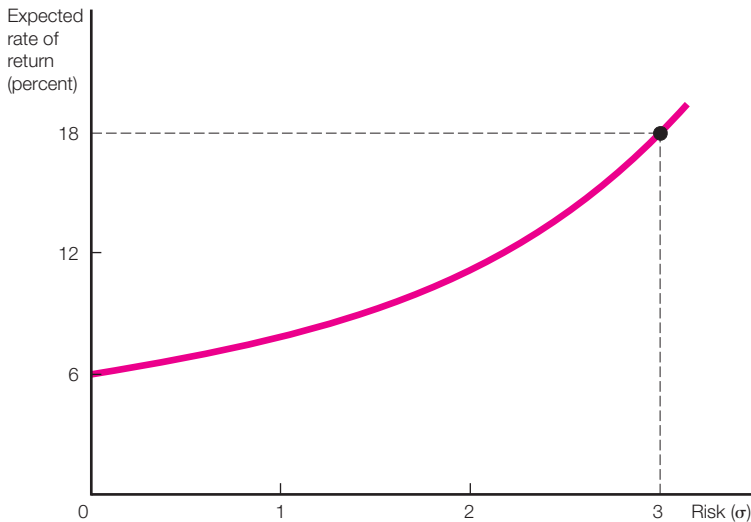
- a. What are the expected present value, standard deviation, and coefficient of variation of investment X?

- b. What are the expected present value, standard deviation, and coefficient of variation of investment Y?
- c. Which investment is riskier?
- d. The president of the Martin Company has the utility function

$$U = 10 + 4P - 0.2P^2$$

where U is utility and P is net present value. Which investment should she choose?

2. William J. Bryan is the general manager of an electrical equipment plant. He must decide whether to install a number of assembly robots in his plant. This investment would be risky because both management and the workforce have no real experience with the introduction or operation of such robots. His indifference curve between expected rate of return and risk is as shown in the figure.
 - a. If the riskiness (σ) of this investment equals 3, what risk premium does he require?
 - b. What is the riskless rate of return?
 - c. What is the risk-adjusted discount rate?
 - d. In calculating the present value of future profit from this investment, what interest rate should be used?



3. The Zodiac Company is considering the development of a new type of plastic. Whether the plastic will be successful depends on the outcome of a research project being carried out at a major university. Zodiac's executives have no reliable means of estimating the university research team's probability of

success. Zodiac's gains (or losses), depending on the outcome of the university research project, are as follows:

Action	Outcome of University Research Project	
	Success	Failure
Zodiac develops plastic	\$50 million	−\$8 million
Zodiac does not develop plastic	0	0

On the basis of the information given, can you calculate the expected value of perfect information? Why or why not? (P.S. You may assume that Zodiac is risk-neutral.)

4. The Electro Corporation, which manufactures television sets, has a fixed cost of \$1 million per year. The gross profit from each TV set sold—that is, the price less the average variable cost—is \$20. The expected value of the number of sets the company sells per year is 100,000. The standard deviation of the number of sets sold per year is 10,000.
 - a. What is the expected value of the firm's annual profit?
 - b. What is the standard deviation of the firm's annual profit?
 - c. What is the coefficient of variation of the firm's annual profit?
5. Richard Miller, a Wall Street trader, says he is risk-neutral. Suppose we let 0 be the utility he attaches to \$100,000 and 1 be the utility he attaches to \$200,000. If what he says is true, what is the utility he attaches to (a) \$400,000? (b) \$40,000? (c) −\$20,000?
6. The chief executive officer of a publishing company says she is indifferent between the certainty of receiving \$7,500 and a gamble where there is a 0.5 chance of receiving \$5,000 and a 0.5 chance of receiving \$10,000. Also, she says she is indifferent between the certainty of receiving \$10,000 and a gamble where there is a 0.5 chance of receiving \$7,500 and a 0.5 chance of receiving \$12,500.
 - a. Draw (on a piece of graph paper) four points on the utility function of this publishing executive.
 - b. Does she seem to be a risk averter, a risk lover, or risk-neutral? Explain.
7. The Oahu Trading Company is considering the purchase of a small firm that produces clocks. Oahu's management feels there is a 50–50 chance, if Oahu buys the firm, that it can mold the firm into an effective producer of washing machine parts. If the firm can be transformed in this way, Oahu believes that it will make \$500,000 if it buys the firm; if it cannot be transformed in this way, Oahu believes that it will lose \$400,000.
 - a. Construct a decision tree to represent Oahu's problem.
 - b. What are the decision forks? (Are there more than one?)

- c. What are the chance forks? (Are there more than one?)
 - d. Use the decision tree to solve Oahu's problem. In other words, assuming that the firm wants to maximize the expected profit, should Oahu buy the firm?
 - e. Before Oahu makes a decision concerning the purchase of the firm, Oahu's president learns that if the clock producer cannot be made into an effective producer of washing machine parts, there is a 0.2 probability that it can be resold to a Saudi Arabian syndicate at a profit of \$100,000. (If the firm cannot be resold, Oahu will lose \$400,000.)
 - (1) Does this information alter the decision tree?
 - (2) Can you think of three mutually exclusive outcomes if Oahu buys the firm?
 - (3) What is the probability of each of these outcomes?
 - (4) What is the monetary value to Oahu of each of these outcomes?
 - f. Use your results in part (e) to solve Oahu's problem under this new set of conditions. In other words, on the basis of this new information, should Oahu buy the firm?
 - g. Oahu's executive vice president discovers an error in the estimate of how much Oahu will gain if it buys the clock manufacturer and turns it into an effective producer of washing machine parts.
 - (1) Under the circumstances in part (d), how big would this error have to be to reverse the indicated decision?
 - (2) Under the circumstances in part (e), how big would the error have to be to reverse the indicated decision?
8. The National Aeronautics and Space Administration (NASA) estimated the probability of a crash of the space shuttle to be 1 in 100,000, whereas the probability was in fact closer to about 0.01–0.02. If a decision tree had been used to determine whether to attempt a launch of the shuttle, what difference, if any, would this have made?
9. The *East Chester Tribune* must decide whether to publish an online Sunday edition. The publisher thinks the probability is 0.6 that this Sunday edition would be a success and 0.4 that it would be a failure. If it is a success, she will gain \$100,000. If it is a failure, she will lose \$80,000.
- a. Construct a decision tree corresponding to the problem, and use backward induction to solve the problem. (Assume that the publisher is risk-neutral.)
 - b. List all forks in the decision tree you constructed; then indicate whether each is a decision fork or a chance fork and state why.
10. Roy Lamb has an option on a particular piece of land, and he must decide whether to drill on the land before the expiration of the option or give up his rights. If he drills, he believes that the cost will be \$200,000. If he finds oil,

he expects to receive \$1 million; if he does not find oil, he expects to receive nothing.

- a. Construct a decision tree to represent Lamb's decision.
- b. Can you tell whether he should drill on the basis of the available information? Why or why not?

Lamb believes that the probability of finding oil if he drills on this piece of land is 0.25, and the probability of not finding oil if he drills there is 0.75.

- c. Can you tell whether he should drill on the basis of the available information? Why or why not?
- d. Suppose Lamb can be demonstrated to be a risk lover. Should he drill? Why or why not?
- e. Suppose Lamb is risk-neutral. Should he drill?

EXCEL EXERCISE: EXPECTED UTILITY

Suppose an entrepreneur has an idea she is excited about. She has wealth of \$100,000 and a utility (U) function

$$U = W^{0.5}$$

where W is her wealth.

To bring the idea to market will require an investment of \$100,000. If the idea is successful (she estimates a 50% chance), she will receive her investment of \$100,000 back plus a return of \$200,000; in other words, if the idea is successful, her wealth will be \$300,000. If the idea fails (she estimates a 50% chance), she will lose her investment, that is, her wealth will be \$0.

Part 1: What is the expected value of the idea? What is the expected value of her wealth if she funds her idea? What is her expected utility if she funds her idea? What is the certainty equivalent of funding her idea? Will she fund her idea?

Part 2: Answer the above questions if her initial wealth was \$1,000,000. Will she fund her idea?

Part 3: Suppose her initial wealth was \$100,000. Could she sell her idea to an investor with wealth of \$1,000,000? What would be the maximum price the millionaire would pay? Remember that once owning the idea, the millionaire would have to put up \$100,000 to bring the idea to market. If the entrepreneur sells the idea to the millionaire at the maximum price, what's the utility of the entrepreneur?

Part 4: Suppose no millionaire exists, but there are 1,000 individuals in the same circumstances as our entrepreneur, that is, with initial wealth of \$100,000 and utility functions of $U = W^{0.5}$. Would these individuals be willing to buy a 1/1,000th share of the idea, in other words, to pay for the right to invest \$100 in the idea, given that there's a 50% chance that they will lose that investment but also that there's a 50% chance they will get that \$100 back plus their share of

the returns of the investment ($\$200,000/1,000 = \200)? What is an individual's expected value of his or her share of the idea? What is the individual's expected value of his or her wealth if they invest? What is the expected utility of wealth if he or she invests? What is the certainty equivalent if these individuals invest? If they do invest and pay the maximum price per share ($1/1,000$ th) of the idea, what's the utility of the entrepreneur?

Part 5: Suppose the entrepreneur feels that there's a 60% chance of success and a 40% chance of failure. The rest of the world, however, believes that the probability is 50–50. Would the entrepreneur fully fund her idea? Should she fully fund her idea, or should she retain n of the 1,000 shares and sell the remaining $1,000 - n$ shares in the market at the maximum price per share you determined in Part 4? What's the entrepreneur's expected utility now?

Let's see how a spreadsheet can answer these questions for you.

Call up your spreadsheet. Enter $=0.5*200000+0.5*-100000$ in cell A1. The result will be the expected value of the idea.

Enter $=0.5*(100000+200000)+0.5*(100000-100000)$ in cell A2. The result will be the expected value of the entrepreneur's wealth if she funds her idea.

Enter 100000 in cell A3. This is the entrepreneur's wealth if she doesn't fund her idea.

Enter $=0.5*(100000+200000)^{0.5}+0.5*(100000-100000)^{0.5}$ in cell A4. This is the entrepreneur's expected utility if she funds the idea.

Enter $=100000^{0.5}$ in cell A5. This is the entrepreneur's expected utility if she does not fund her idea. Comparing cell A4 with cell A5 shows you that $A5 > A4$, and so the expected utility from not funding exceeds the expected utility from funding, so that the idea is not brought to market despite the expected wealth increase that occurs by funding the idea; in other words, compare cell A2 with \$100,000.

Enter $=A4^2$ in cell A6. This is the entrepreneur's certainty equivalent if she funds the idea. Comparing this with \$100,000 confirms that she will not fund the idea.

Enter $=0.5*200000+0.5*-100000$ in cell A7. This is the expected value of the idea for the millionaire.

Enter $=0.5*(1000000+200000)+0.5*(1000000-100000)$ in cell A8. This is the expected value of the millionaire's wealth if he invests in the idea.

Enter $=0.5*(1000000+200000)^{0.5}+0.5*(1000000-100000)^{0.5}$ in cell A9. This is the expected utility of the millionaire's wealth if he funds the idea.

Enter $=1000000^{0.5}$ in cell A10. This is the millionaire's expected utility if he does not fund the idea. Note that the value in cell A9 exceeds the value in cell A10 so the millionaire funds the idea.

Enter $=A9^2$ in cell A11. This is the certainty equivalent of the millionaire funding the idea. This is greater than \$1,000,000 (the certainty equivalent of not funding) confirming that the millionaire would fund the investment, but the

original entrepreneur would not. Can you see from this why hedge fund operators seek out rich investors? They do so because rich investors are willing to put up cash that small investors are unwilling to put up.

Enter $=0.5*(1000000+200000-A13)^{0.5}+0.5*(1000000-100000-A13)^{0.5}-A10$ in cell A12. Then substitute values of your choice in cell A13 until the value of cell A12 becomes 0. You should start with a value around the certainty equivalent minus 1,000,000.

The value in cell A13 that gives a 0 in cell A12 is the maximum amount the millionaire would pay the entrepreneur for the rights to the idea. Note it is the amount that makes the entrepreneur indifferent to the choice between purchasing the idea and then investing in the idea or just sitting on their \$1,000,000. This value is approximately \$44,375.

Enter $=100000+A13$ in cell A14. This represents the amount of wealth of the entrepreneur if she sells the rights to the millionaire at the maximum price the millionaire is willing to pay.

Enter $=A14^{0.5}$ in cell A15. This represents the entrepreneur's expected utility if she sells her idea to the millionaire at the millionaire's maximum price.

Enter $=0.5*200+0.5*-100$ in cell A16. This represents a single shareholder's (out of 1,000 shareholders) expected value of the idea.

Enter $=0.5*(100000+200)+0.5*(100000-100)$ in cell A17. This represents the expected wealth of any individual shareholder.

Enter $=0.5*(100000+200)^{0.5}+0.5*(100000-100)^{0.5}$ in cell A18. This represents the expected utility of any individual shareholder.

Enter $=100000^{0.5}$ in cell A19. This represents the expected utility of an individual if he or she fails to become a shareholder. You will note that the value of cell A18 exceeds the value of cell A19. Thus, the individuals wish to become shareholders.

Enter $=A18^2$ in cell A20. This is the certainty equivalent of a shareholder. This exceeds \$100,000, confirming that the individual should become a shareholder.

Enter $=0.5*(100000+200-A22)^{0.5}+0.5*(100000-100-A22)^{0.5}-A19$ in cell A21. Then substitute values of your choice in cell A22 until the value of cell A21 becomes 0. You should start with a value around the certainty equivalent minus 100,000.

The value in cell A22 that gives a 0 in cell A21 is the maximum amount the individual would pay the entrepreneur for the rights to 1/1,000th of the idea, that is, a share. Note that it is the amount that makes the individual indifferent between purchasing the 1/1000th of the idea and then investing \$100 in the idea and just sitting on their \$100,000. This value is approximately \$49,944.

Enter $=100000+1000*A22$ in cell A23. This represents the amount of wealth of the entrepreneur if she sells the rights to the idea to 1,000 individuals at the maximum price each individual is willing to pay.

Enter $=A23^{0.5}$ in cell A24. This represents the entrepreneur's expected utility if she sells 1,000 shares of her idea to the 1,000 individuals at each individual's maximum price.

Why is it better to sell 1,000 single shares to 1,000 individuals rather than 1,000 shares to one individual? Because the millionaire faces more risk than the single individual.

Subtract A11 from A8, that is, enter $=A8-A11$ in cell A25. Cell A25 represents the risk premium the millionaire must receive to purchase the idea. Subtract A20 from A17, that is, enter $=A17-A20$ in cell A26. Cell A26 represents the risk premium that each single investor would have to receive to invest in a single share of the idea. Because the investors are investing so little (\$100), little is at risk, and they are basically risk averse. You will notice that their certainty equivalent (cell A20) is virtually the same as $\$100,000+A22$; that is, their initial wealth plus what they are willing to pay for a share.

Now suppose that the entrepreneur feels that there is a 60% chance of success of the idea. This could be because the entrepreneur has more intimate knowledge about the idea than the market (which views the probability of success at 50%).

Enter $=0.6*(100000+200000)^{0.5}+0.4*(100000-100000)^{0.5}$ in cell A27. This is the expected utility if the entrepreneur fully funds her idea.

Enter $=A27^2$ in cell A28. This is the certainty equivalent of the entrepreneur fully funding her idea.

Comparing A27 with A5 and A28 with \$100,000 shows that the entrepreneur will now be willing to fully fund her idea. But should she? Is there a way for her to get more utility? You already know the answer. The utility in A24 exceeds the utility in cell A27 so she should sell her idea in 1/1,000ths to 1,000 individuals. But can she do even better by retaining some (n) of those 1,000 shares? Let's see.

Enter $=0.6*(100000+200*A30+(1000-A30)*A22)^{0.5}+0.4*(100000-100*A30+(1000-A30)*A22)^{0.5}$ in A29. This is the expected utility of the entrepreneur if she retains A30 shares and sells 1,000-A30 shares. Can you find a value of A29 that exceeds the value in A24? In other words, can the entrepreneur get a higher expected utility from selling 1,000-A30 shares than from selling all 1,000 shares?

The answer is yes. You can experiment by substituting various values in cell A30 (or you can learn to use Excel's solver, which will solve for the A30 that maximizes the value of cell A29). Either way you will find that at about 385, A30 maximizes the entrepreneur's expected utility. When all is said and done, the ability to securitize a project enables an idea with positive expected value to be brought to market whereas if left to the initial entrepreneur, it would not see the light of day.

Enter $=A29^2$ in cell A31. This is the certainty equivalent of the entrepreneur if she securitizes her idea but retains 385 shares. The entrepreneur has now parlayed her initial wealth of \$100,000 into a certainty equivalent value of almost \$156,000.

Where does the entrepreneur find the 1,000 investors? Probably through an investment banker. Thus, some of the profit from securitizing will be taken by the investment banker for finding the willing investors. Does it seem reasonable that in a world with hundreds of millions of investors, that 1,000 could be found that look like our entrepreneur? Very likely. There's another bonus here. Investors like it when the initial entrepreneur retains shares. This means that the entrepreneur does not take the money and run. The entrepreneur has "skin in the game" and is very likely to be hands-on and work diligently to make sure the idea succeeds.

We can show how easy it is to use Excel Solver to find the optimal number of shares in cell A30.

First, duplicate the formula used in cell A29 and in cell A33. This is

$$=0.6*(100000-200*A34+(1000-A34)*A22)^{0.5} \\ +0.4*(100000-100*A34+(1000-A34)*A22)^{0.5}$$

Next, in the toolbar in the top row of Excel, click on "Tools." A menu will appear. Click on "Solver." A solver parameter box will appear. In the "Set Target" box, enter A33. In the "Equal" box, click on "Max." In the "By Changing Variable Cells" box, enter A34. Then click on the "Solve" box. A "Solver Results" box will appear to announce a solution is found. Click the "OK" box.

A maximum utility value of 394.92 will appear in cell A33. The number of shares the entrepreneur should self hold (385.163) will appear in cell A34.

PRINCIPAL-AGENT ISSUES AND MANAGERIAL COMPENSATION

PRINCIPAL-AGENT ISSUES

How would you behave if we gave you millions of dollars and tickets to Las Vegas? This same question is faced by most upper-level managers. We assume managers seek to maximize value for shareholders; for the most part this is true. But managers face situations where their personal utility function conflicts with that of being an agent of the firm. This is at the heart of the **principal-agent issues** we discuss in this chapter. A conflict of interest arises when one person, the agent, makes decisions on behalf of another, the principal, in the face of uncertainty. The uncertainty is caused by imperfect monitoring of managerial behavior. In publicly held firms, managers (agents) make decisions that affect the wealth of shareholders (principals).

When the interests of the principal and agent are identical, we do not worry about this issue. For example, the captain and sailors on a ship in a storm need each other to save the ship. Because the sailors know the captain wishes to save his own life, the sailors may be confident he will make decisions that are also in the best interests of the sailors. Issues arise when interests and preferences are not identical. Hence principal-agent issues are not confined to business; we see them arise in most professions. For example, will a general, who is not on the battlefield, devise strategies that best serve the soldiers under his command? Generals may wish to win battles; soldiers prefer to stay alive. Can we be sure that elected

LEARNING OBJECTIVES

Principal-Agent Issues

The Diverging Paths of Owners and Managers

The Principal-Agent Situation

The Effect of Risk, Information, and Compensation on Principal-Agent Issues

Resolving the Incentive Conflict When Output Is Risky and Effort Is Not Observable

Some Refinements to Managerial Compensation

Principal-Agent Issues in Other Contexts

Product Liability and the Safety of Consumer Goods

Summary

Problems

Excel Exercise: Moral Hazard

Principal-agent issues When managers (agents) make decisions that affect the wealth of shareholders (principals).

politicians really serve their constituents and are not nudged away from that goal by interest groups? Will a plaintiff's lawyer be tempted to advise her client to go to trial (which enhances the visibility of the lawyer) rather than to accept a good settlement offer? Will a doctor prescribe treatment that is best for his patient, or might the doctor use the patient to further a research agenda or prescribe treatment that leads to the highest payment for the doctor? It is human nature to face such conflicts.

To understand the principal-agent relationship in business, we need to consider the effects of uncertainty and information. One form of uncertainty occurs because the outcomes of agents' actions are not linked in a totally deterministic way with their effort. Knowledge of results does not necessarily imply anything about effort. This lack of a direct link is caused by information asymmetry; that is, agents and principals do not share common sets of information.

The traditional corporate governance structure is straightforward. Shareholders own the firm's assets and assume the risks of doing business. Any residual profit from the actions of managers is split among the shareholders. Shareholders hire managers (agents) to run the business. Agents choose actions from a number of alternative possibilities (such as allocating resources within the firm). Actions affect the welfare of both the agent and the principals. However, the interests of the two parties are not necessarily identical.

Principals and agents play a noncooperative game. Principals determine rules that assign agent compensation as a function of the principal's observation of firm performance. But there is asymmetric information here: Agents (managers) have more information about the action than does the principal. The agent's action is not directly observable by the principal, and the outcome of the action is not completely determined by the agent's action.

The first asymmetric information issue is the hidden action or moral hazard. The term moral hazard arises because of insurance considerations. For example, obtaining fire insurance dulls incentives for caution and can even create incentives for arson. The most common hidden action issue in the corporate world is determining the effort of agents. Effort has a disutility to the agent but has a value to the principal because it increases the probability of a favorable outcome.

In this chapter we discuss the principal-agent issues common to most managers. Our guidelines will help managers anticipate these issues and minimize their disruptive effect. Many principal-agent issues arise under the umbrella of the firm because of the inherent conflict between the collective and individual. We show how incentive schemes like bonuses, equity, and options help converge preferences and resolve much of this conflict.

We also show why managers need to anticipate and control the incentive conflicts that arise between shareholders and creditors. For example, such conflicts might prevent managers from adopting investment projects despite the fact that they add value.

Finally, we explore the implications of principal–agent issues in looking at the evolution of product liability laws. In particular, these laws provide incentives for managers to produce safe products. However, we also show that in the information age, product liability laws may not be necessary, and the price mechanism may provide similar protections for consumers.

THE DIVERGING PATHS OF OWNERS AND MANAGERS

One of the most important principal–agent issues encountered in business is that between the owners of a firm and its managers. The owners typically are shareholders who purchase the stock as an investment, investors who simply buy shares in a mutual fund, or pensioners whose assets are invested in many companies. Most investors probably are interested in maximizing the value of their investment, which means either maximizing the income their assets yield or maximizing the value of those assets.

Clearly shareholders are concerned about the value of their shares. One question we can ask is whether managers care as much as shareholders do about maximizing share value. And if they don't, what other goals might managers have? Let us make a plausible, though not necessarily complete, list of alternative goals that may prevent managers from always taking actions to maximize firm value:

- *Minimizing effort:* Increasing profit often takes hard work. There is always disutility to work given the opportunity cost of leisure. So managers face the following question: Will I be rewarded for my hard work? Given a choice between two activities, one involving a little more effort and the other a lot more effort, which will most managers choose? The manager supplies the effort while the owner reaps the profit. Managerial behavior is largely driven by how owners structure compensation. Some examples suggest that many marginal decisions of managers tilt in favor of less effort rather than higher profit.
- *Maximizing job security:* Many decisions of managers involve risk. Often risky projects are characterized by a high potential reward or a large potential loss. Managers may be disinclined to make risky choices that could jeopardize their employment. Suppose an investment decision carries a large probability of very high returns and a small chance of failure that would cause the firm itself to fail. Shareholders, being diversified, might be inclined to accept this risk. Managers, however, might be more concerned with the downside (they do not share the upside profit and could lose their jobs given the downside risk), so they might be tempted to avoid such risky choices.
- *Avoiding failure:* Managers can be rewarded for good performance and penalized for bad results. If a risky project is undertaken, the manager is rewarded if results (due somewhat to chance) are favorable and penalized if results are unfavorable.

Often managers believe that bad results are much more likely to be noticed than good results. If so, they are disinclined to take risks.

- *Enhancing reputation and employment opportunities:* Although we argue that sometimes reputation is promoted by doing things that benefit shareholders, this is not always so. For example, a CEO with ambition to hold public office might be more concerned about showing himself to be a “good citizen” than maximizing the profit of the firm. Therefore, prices charged might be lowered below their profit-maximizing level. Alternatively, a manager might conduct contract negotiations with another firm partly with a view to establishing a personal relationship that could be a springboard to a new job.
- *Consuming perquisites:* Examples can include luxury travel, expensive artwork in the office, corporate donations to favorite charities, and employing favored people.
- *Pay:* The manager presumably works for pay; and as we shall see, both the level and structure of the compensation package become important parts of the principal–agent story.

THE PRINCIPAL–AGENT SITUATION

Figure 15.1 shows a diagram of the principal–agent situation. The principal employs an agent who performs a task that results in a benefit to the principal. The benefit is called output. The principal must pay the agent. This compensation can be a fixed sum or it can depend on the output. The level of output depends on the quantity and quality of the effort provided by the agent. If this effort were observable by the principal, the principal could simply require a certain level of effort, verify that this level of effort was provided, and compensate the agent accordingly. But effort is not always observable or measurable. The principal–agent environment is one in which effort cannot be perfectly monitored by the principal and therefore cannot be directly rewarded. The efficient solution requires some alignment of interests of the two parties. In this way, even if the principals cannot perfectly see what the agents are doing, they still can be assured that what is good for the agent is also good for the principal.

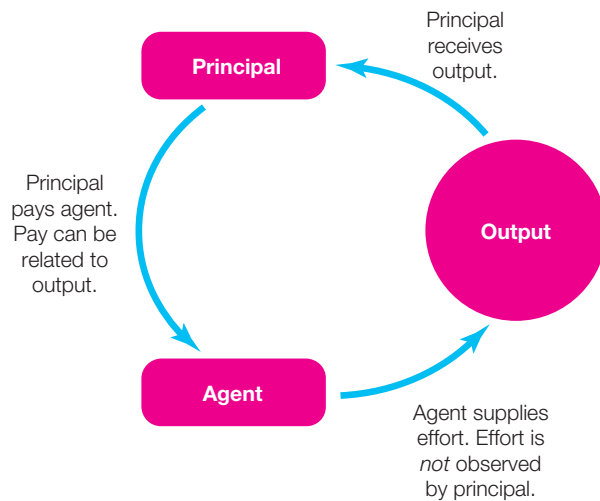
Examples of Principal–Agent Issues

Returning to a more familiar theme in this book, we believe managers generally take actions to maximize firm value. This seems reasonable because the firm’s owners (shareholders) are clearly interested in the value of their shares, and this value rests on the long-term profitability of the firm. We would expect the shareholders to appoint a board of directors and management team that pursue the profit objective. However, corporate governance is more complex. A subject of considerable research is whether boards really do control managers or vice versa. What is clear is that, although managers are accountable in the long run to the

FIGURE 15.1

The Principal-Agent Problem

The principal employs an agent to undertake a task (produce output for the principal). The agent must expend effort to produce output; the more effort, the more output. Because the principal cannot observe (and therefore cannot reward) the effort, the agent tends to “shirk,” or reduce effort, which in turn reduces the output for the principal.



owners, they still have considerable discretion in making decisions. Consider some examples.

1. Managers at a life insurance company are thinking about selling a new product through one of two distribution channels. One alternative is to use an existing distribution channel: a network of independent agents who are paid commissions. This approach is safe but has a low expected profit. The other choice is to establish an electronic distribution system. This is new and untried; but if it succeeds, profit could be very high. If the expected profit is high enough for the e-business channel, this system might be best for the shareholders. Shareholders would certainly benefit from the higher expected profit. Moreover, they probably are not too concerned about risk because most shareholders hold a number of securities and can diversify their risk. However, managers might worry more about risk because if the e-project is chosen and the firm has bad luck, the managers might lose their jobs. Therefore, managers might be tempted to play it safe with the existing distribution network even though electronic distribution could be better for shareholders.

2. Managers at many firms make charitable contributions, and it can be argued that some level of charitable giving benefits shareholders. This visibility

might stimulate demand for the firm's products and cast the firm in a favorable light with legislators and regulators. Charity also might bring some benefit in courts. For example, a car manufacturer that donates considerable sums to university safety research might find that this charitable giving is an important signal of its commitment to safety, which helps in the defense of lawsuits. Now consider a manager who has a personal agenda: He is a strong supporter of environmental causes. This manager not only increases the level of charitable giving above what is optimal for the shareholders but also redirects it to his favored environmental charity.

THE EFFECT OF RISK, INFORMATION, AND COMPENSATION ON PRINCIPAL-AGENT ISSUES

Managerial Behavior and Effort

We cannot understand the world of managers until we discuss the costs of effort. Achieving a target level of profit requires that managers incur some personal cost, which we call effort. This effort might simply be the manager's time required to attain that level of output, and the cost to the manager is the value of that time. But time is not the only dimension. Certain tasks require less pleasant work than others. Spending time with clients on the golf course may be valued differently by a manager than spending time bargaining with the union over wages. Managers may sacrifice other things to attain a profit goal. For example, to increase expected profit, a manager may cut down on perquisites and employ the best engineer rather than giving the job to her favorite niece. The cost to the manager may be a direct cost or an opportunity cost. The personal cost to the manager of making a decision reflects the quantity and quality of effort required. Given a choice among activities, most managers will exert some bias toward tasks that require less effort. Of course shareholders prefer managers to exert more effort.

We show how managers can work with owners to resolve some of these principal-agent issues. To do so, we solve profit-maximizing equations similar to those in earlier chapters. However, in earlier chapters we saw how managers often try to choose a quantity (and a price) for a given product to maximize profit, or they purchase the least costly input mix. In the principal-agent context, owners can design the manager's compensation to maximize profit; that is, owners can modify managerial behavior via compensation structures. In doing this, they recognize that how managers are compensated affects the managers' choice of effort.

We now represent the situation more precisely. The profit to the firm, π , depends on the manager's effort, e . We write this function as $\pi(e)$ to remind us of the impact of effort. For the moment we assume profit is not risky. Once the manager chooses effort, we can forecast the profit with certainty. Profit is total revenue minus total costs, and the manager's effort affects profit by changing the

total revenue. Therefore we write revenue as $R(e)$ to show that it depends on effort. Finally we divide total cost between the manager's compensation, S , and all other costs, C . Profit is now written as

$$\begin{aligned}\text{Profit} &= \{\text{Revenue}\} - \{\text{Costs}\} \\ &= \{\text{Revenue}\} - \{\text{Managerial compensation} + \text{Other costs}\} \\ \pi(e) &= \{R(e)\} - \{S + C\}\end{aligned}$$

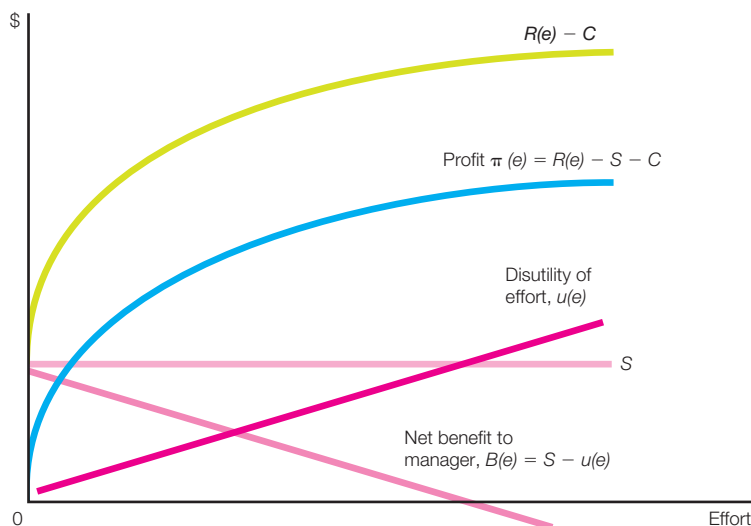
More effort on the part of the manager results in higher revenue. However, for the moment we assume that S is a flat salary and cannot change relative to effort. For simplification, we assume other costs are not affected by the manager's effort. We could assume that managers could reduce costs if they tried harder. But this would complicate our analysis, and the main ideas we develop will be substantially the same anyway.

Prior to deducting the manager's salary, let $\pi(e) = R(e) - C$; and profit after deduction of salary, $\pi(e) = R(e) - S - C$, has the upward slope shown in Figure 15.2. If the owners were to choose the level of effort they wish the manager to supply, they would choose the highest feasible level. However, owners cannot simply choose the level of effort, e , to maximize profit. The manager, not the owners, must choose e . The manager has his or her own objectives and chooses a level

FIGURE 15.2

The Principal-Agent Problem with a Flat Salary

Because the net benefit to the agent is the salary minus the disutility of effort, this benefit declines as effort increases. Accordingly, the agent reduces effort, which reduces revenue.



of effort to maximize these personal objectives. The manager is not only a member of the collective; she is also self-interested.

The objective of the manager is to maximize the net benefit of employment. The manager obtains income from employment and prefers this income to be high. To gain this income, the manager must supply effort. The cost to the manager of supplying effort is shown as the function $u(e)$; the net benefit to the manager is her compensation minus the cost in effort it took to achieve it. So there is a **disutility of effort**, which is a measure of the cost to the manager of supplying effort. At the very least, the manager must suffer an opportunity cost. More effort involves more cost, or disutility, to the manager, so $u(e)$ slopes upward. The net benefit to the manager of working at a given level of effort is now shown as

$$B(e) = S - u(e)$$

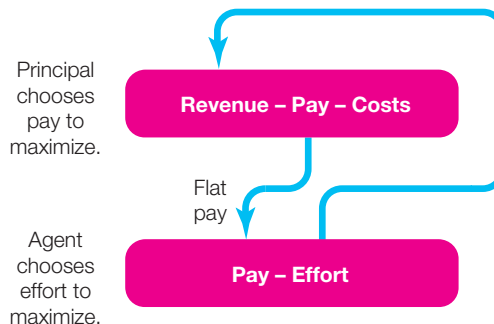
Because the salary is constant and disutility increases with effort, the net benefit, $B(e)$, must slope downward, as shown in Figure 15.2. Therefore, the manager, who bears all of the cost of effort but gets none of the reward, is better off with as little effort as possible—in this case, zero effort. In contrast, the shareholders, who get all of the benefit of the manager's effort but pay none of the cost, want maximum effort from the manager.

The incentives for managers and owners when managers are paid a straight salary are summarized in Figure 15.3. At the top of the diagram, the principal wants to maximize profit but cannot control the manager's effort. The owner pays a flat salary. At the bottom of the picture, the manager wants to maximize the net benefit of employment; but because pay is constant relative to effort, maximizing this benefit is achieved only by reducing effort. Managers tend to exert less effort when effort is not correlated to a higher net benefit.

FIGURE 15.3

The Principal-Agent Problem with Flat Pay

The owner pays the manager a flat salary. The manager chooses to minimize effort; therefore revenue and profit fall.



STRATEGY SESSION: Private Equity Firms Deal with the Principal-Agent Problem

Agency problems have many faces. The issue in corporate governance is whether managers perform the way the owners of the firms wish. The problem arises because the owners (the principals, that is, the shareholders) can't directly monitor the actions of their agents (top corporate management). Monitoring by principals is difficult because shareholders have day jobs and don't have the time (or skills) for monitoring. When private equity firms take over a company, the many principals become a few whose responsibility is to monitor the company. If managers aren't behaving in the interests of the owners, lost profit comes out of the owners' pockets (as it does for a publically owned company). The difference is now that the previously hidden actions of the agents are no longer hidden from the new principals.

What's the first thing to go? Often it's the corporate jet. A corporate jet costs an average of \$1 million

per year to operate. While such jets provide more security for top executives, cut down on delay and transfer times, allow day trips to less urbanized areas, and so on, they have come under the scrutiny of the new principals. Other items hitting the chopping block are golf club memberships, retreats to pricey resorts, and corporate-owned houses/apartments. When TXU Corporation was purchased by private equity firms TPG and Kohlberg Kravis Roberts & Co. (KKR & Co.), the Gulfstream jet went. When KKR & Co. purchased RJR Nabisco, seven of the eight corporate jets and more than twelve company-owned houses/apartments were sold off.

Source: Jason Kelly, "Corporate Jets Often First Thing to Go After Leveraged Buyouts," *Bloomberg News*, April 8, 2011, at www.bloomberg.com/news/2011-04-08/corporate-jets-often-first-thing-to-go-after-leveraged-buyouts.html.

Resolving the Incentive Conflict If Effort Is Observable

How can owners motivate managers to work harder? If owners can observe the effort provided by managers, the owners can reward managers directly in relation to the effort. Look at Figure 15.4. This is similar to Figure 15.3 except for the inclusion of incentive pay. The pay by the owners is scaled to the manager's effort. Now the manager's choice of effort is more complex. Although effort is unattractive in itself, it has the compensating advantage of increasing the manager's pay; so the manager is now persuaded to increase effort, which in turn increases revenue and profit.

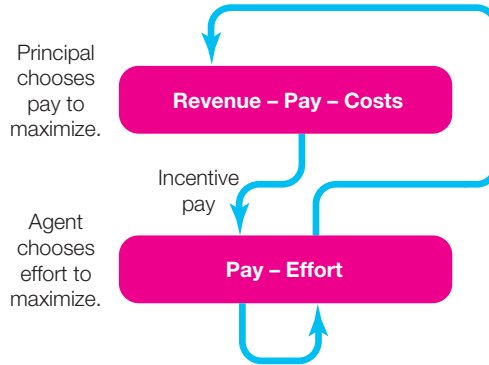
To see how this works in more detail, let the owners structure the compensation in two parts. The first part, K , is a fixed amount. $U(e)$ is an additional amount that varies with managerial effort. So $U(e)$ represents the reward paid by the firm for managerial effort, while the lowercase $u(e)$ represents the monetary value of the disutility of effort to the manager. When $U(e) = u(e)$, the manager is fully compensated for effort, as we will see shortly. Because compensation now is a function of effort, we write it as $S(e)$

$$S(e) = K + U(e)$$

FIGURE 15.4

Motivating Managers When Effort Is Observable

The owner pays incentive compensation based directly on effort. Because the agent's pay increases with effort, the agent increases his effort. This in turn increases the firm's revenue. If the increase in revenue exceeds the increase in pay, profit increases.



This is shown in Figure 15.5. The profit now is

$$\begin{aligned}\pi(e) &= R(e) - S(e) - C \\ &= R(e) - [K + U(e)] - C\end{aligned}$$

With full information regarding the actions of managers, owners can easily identify their preferred effort level as that which maximizes profit. Note that only $R(e)$ and $U(e)$ depend on effort, so the maximum is obtained by

$$\frac{\Delta \pi(e)}{\Delta e} = \frac{\Delta R(e)}{\Delta e} - \frac{\Delta U(e)}{\Delta e} = 0$$

which simply says the marginal benefit from effort (in terms of increased revenue, $R(e)$) must equal the marginal cost of compensating the managers for effort. Note that profit achieves a clear maximum in Figure 15.5 at effort level e^* . This is because profit reflects both the benefits to the shareholders and the costs to the manager for which the owners now have to pay him. The shareholders would wish to maximize profit at effort level e^* .

QUANT OPTION

Managers will maximize profit when

$$d\pi(e)/de = dR(e)/de - dU(e)/de = 0$$

PROBLEM SOLVED: Setting Optimal Compensation When Output Is Not Risky and Effort Is Not Observable

The weekly revenue, $R(e)$, of a retail store called Sporting Goods depends on the manager's effort, e ; the more effort she provides, the higher the revenue. The effort is the number of hours worked. However, because many of the working hours are not spent in the store, the owner does not know directly how many hours the manager actually works. Assume that even if the manager supplies no effort, revenue would equal 3,500. Effort by the manager causes revenue to increase by $100e^{0.5}$. Thus revenue can be written as

$$R(e) = 3,500 + 100e^{0.5}$$

But the manager incurs a disutility from supplying the effort of

$$u(e) = 853.55 + (7.07)e$$

In addition to paying the manager's salary, the shareholders must pay production costs of 1,000. To be persuaded to work for the store, the manager must

receive a net benefit, $B(e)$, of 1,000. If she receives less, she will leave and take another job.

- How much effort, e , will the manager supply?
- How much weekly profit, $\pi(e)$, will be made?
- How much of this profit will be paid to the manager?

First, note that the owners of Sporting Goods can infer the level of effort provided by the manager from the week's revenue. This can be done by solving the revenue equation backward

$$\begin{aligned} R(e) &= 3,500 + 100e^{0.5} \\ \frac{R(e) - 3,500}{100} &= e^{0.5} \\ e &= \left(\frac{R(e) - 3,500}{100} \right)^2 \end{aligned}$$

The owners can pay the manager a compensation package $U(e)$ that includes a component to com-

QUANT OPTION

$$\text{For } R(e) = 3,500 + 100e^{0.5},$$

$$\Delta R(e)/\Delta e = 0.5(100)e^{-0.5} = 50e^{-0.5}$$

$$\text{For } u(e) = 853.55 + 7.07e,$$

$$\Delta u(e)/\Delta e = 7.07$$

However, this ignores the main issue. Managers, by the nature of their duty, have better information with which to make decisions than the shareholders. Managers are appointed precisely because they have the time and skill to make decisions. Shareholders have no time to manage every firm in which they own

pensate indirectly for her effort and a bonus $\alpha\pi(e)$ related directly to profit. The profit, before payment of the bonus, is now

$$\pi(e) = R(e) - U(e) - C = R(e) - u(e) - C$$

If the compensation component $U(e)$ is enough to compensate the manager for the disutility of her effort, $U(e) = u(e)$, she will choose to maximize her bonus, which is the same as maximizing profit because the bonus is a proportion of profit.

$$\frac{\Delta\pi(e)}{\Delta e} = \frac{\Delta R(e)}{\Delta e} - \frac{\Delta u(e)}{\Delta e} = 0$$

The solution to this type of problem is now familiar. The manager simply equates the marginal benefit and marginal cost

$$\frac{\Delta R(e)}{\Delta e} = 0.5(100)e^{-0.5} = \frac{\Delta u(e)}{\Delta e} = 7.07$$

Setting marginal benefit equal to marginal cost yields

$$0.5(100)(e)^{-0.5} = 7.07$$

so

$$\begin{aligned} e^{0.5} &= 50/7.07 \\ e &= 50^2/7.07^2 \\ &= 2,500/50 \\ &= 50 \end{aligned}$$

The manager works 50 hours per week. We can plug this effort level into the profit equation to give weekly profit

$$\begin{aligned} \pi(e) &= R(e) - u(e) - C \\ &= [3,500 + 100e^{0.5}] - [853.55 + (7.07)e] - 1,000 \\ &= [3,500 + 100(50^{0.5})] - [853.55 + (7.07)(50)] - 1,000 \\ &= [3,500 + 707.11] - [853.55 + 353.55] - 1,000 \\ &= 2,000 \end{aligned}$$

If the manager is to receive a net benefit, $B(e)$, of 1,000, clearly she must receive 50% of the profit, so $\alpha = 0.5$.

some shares; nor do they have the expertise to evaluate each decision. Nor can a board of directors manage every decision. Managers are generally given considerable freedom in operating discretion. In the long run, the impact of their stewardship of a firm can be inferred from how it performs. But the long run can be very long, and this permits managers considerable flexibility in balancing their own ambitions against those of the firm's owners. This discretion means that owners do not observe all the actions of the manager. In effect, it is too costly for owners to perfectly monitor managerial effort.

Resolving the Incentive Conflict If Effort Is Not Observable: Incentive Compatibility

The nature of most principal-agent issues is that when effort is not observable by the principal, it is difficult to reward or penalize directly. The inability of owners to completely observe managerial actions gives managers some degree of freedom to pursue their own objectives, instead of always acting in the best interest of the principal.

Let us assume the firm's revenue, $R(e)$, is riskless and determined solely by the effort of the manager. The shareholders can ask, What level of effort is necessary to produce this level of revenue? In this way shareholders deduce the level of effort by observing revenue. If $R(e)$ slopes upward, as shown in Figures 15.2 and 15.5, shareholders can read backward on the graph from a given R on the vertical axis to an implied level of e on the horizontal axis. For example, if shareholders observe revenue level $R(e^*) = C$ in Figure 15.5, then the effort level necessary to achieve this metric must be e^* .

QUANT OPTION

Managers will maximize profit if

$$d\pi(e)/de = dR(e)/de - du(e)/de = 0$$

Because shareholders can infer the level of effort, observing R is as good as observing e . We can now calculate what e must have been, so we can make the same compensation payment $S(e) = K + U(e)$ as before. However, there is one snag. This ensures that the manager gets a retrospective payment that fully compensates her for effort already provided. Anticipating this reward for effort, the manager is induced to work harder. But how can the shareholders ensure the manager will choose exactly the level of effort, e^* , that they prefer?

Shareholders could make a simple but important change to the manager's compensation: Give managers a share, α , of profits, $\pi(e)$. The new compensation is $S(e) = U(e) + \alpha\pi(e)$. They replace the fixed amount K with a profit-sharing bonus.¹ Naturally, the manager does not obtain the bonus until the end of the period. Therefore, owners choose a level of bonus so the overall package is competitive and attracts and retains skilled managers. We divide the manager's compensation into two parts

$$\begin{aligned} S(e) &= (\text{Salary}) + (\text{Bonus}) \\ &= [U(e)] + [\alpha\pi(e)] \end{aligned}$$

$\alpha\pi(e)$ is a share of profit after the salary is paid. So the profit is now

$$\pi(e) = R(e) - U(e) - C$$

This makes the bonus a real equity share in the firm.

The net benefit to the manager is now

$$\begin{aligned} B(e) &= S(e) - u(e) \\ &= U(e) + \alpha\pi(e) - u(e) \\ &= \alpha\pi(e) \end{aligned}$$

1. Note that we have set part of the compensation, $U(e)$, in relation to effort. Although effort is not directly observed, the principal can infer effort from the actual level of profit.

The last step was taken by setting the monetary compensation for effort, $U(e)$, at a level sufficient to offset the disutility of effort, $u(e)$. We have a clear alignment of the interests of the managers, who get a net benefit of $\alpha\pi(e)$, and the owners, who get the remainder $(1 - \alpha)[\pi(e)]$. Now both are interested in maximizing the net profit, $\pi(e)$.

By giving managers a share of the profits, owners align their preferences with those of managers. Consider what level of effort owners prefer relative to what level the managers will choose. The owners get a portion of the profit, $(1 - \alpha)\pi(e)$, and the manager gets $\alpha\pi(e)$; so whatever the level of α , both shareholders and managers are happiest if $\pi(e)$ is maximized. We can think of this process in two stages:

1. The manager chooses a level of effort to maximize $\pi(e)$. This is achieved where the marginal benefit of effort equals the marginal disutility of the cost of effort.
2. The firm's owners choose a level of α such that the compensation package is competitive.

We have now introduced an important concept: **incentive compatibility**. Because the agent and the owners share in the profit of the firm, their incentives are aligned and compatible. We refer to contracts that have this alignment of interests as incentive-compatible contracts.

Incentive compatibility When the agent and the owners share in the profit of the firm; and the agent's effort maximizes the principal's profit.

RESOLVING THE INCENTIVE CONFLICT WHEN OUTPUT IS RISKY AND EFFORT IS NOT OBSERVABLE

The incentive issues between owners and managers change when revenue is risky and effort is not observable. With no risk, owners can infer the level of managerial effort from the firm's profit. However, when profit is risky, owners are rarely certain whether high profit is due to high effort or simply good luck (a strong economy) and whether low profit is due to low effort or bad luck. Poor management can occasionally result in short-term high profit due to random events. Similarly, determined effort can sometimes come unstuck due to the vagaries of the market.

Risk Sharing

When output is risky and effort is not observable, owners need to structure managerial compensation around two competing ideas: risk sharing and efficiency. The efficiency idea has been dealt with previously. The manager is inclined to attend to his or her own interests, so it is desirable for the firm's owners to align incentives with profit-sharing bonuses, equity participation, stock options, and similar instruments. However, the firm's profit and its equity value are uncertain. Moreover, this volatility is partly outside the control of the managers and owners. Profit

and share price are affected by macroeconomic factors such as changes in interest rates, employment, inflation, foreign exchange rates, and movements in stock market indexes. Thus bonus plans impose some risk on managers.

Owners can also design managerial compensation based on who can tolerate risk at the least cost, or who is the least risk-averse. Owners tolerate risk in that managers generate some value to be divided among the stakeholders, and this value is inherently risky. The value can be expressed as the periodic profit or the value of the equity. As residual claimants, the shareholders normally get the risky profit or equity, but first they must pay the managers. Should the shareholders take out a fixed sum (a flat salary) and pay this to the managers, or should they simply give the managers a share in the risky profit or equity? For the shareholders, the riskiness of profit may not be too problematic. Most shares are owned by investors or institutions that are quite diversified. Many individual investors hold several assets in their portfolios and are concerned not about the riskiness of each stock, but about the risk in their whole portfolio. Unless the stocks are highly correlated, the risk in the portfolio can be quite modest. Therefore, investors, being able to diversify, can tolerate the risk in an individual stock. If, for example, the risk in one stock in her portfolio increased, an investor could offset this by spreading her capital over a few more stocks. Institutional investors are often much better diversified than individual investors and can easily tolerate the risk of each individual stock.

In contrast, managers are usually much less diversified. For the typical manager, the compensation received from an employer, and the equity stake in that employer, is a large proportion of his or her total wealth. Fluctuation in the value of bonuses or stock options can have a big impact on the manager's net worth. Hence we expect managers to be quite averse to risk in their compensation plan. This does not mean they will not accept a risky compensation plan. Rather, they need to be compensated for the risk; they need a **risk premium**.

Risk premium The minimum difference a manager requires to be willing to take a risk.

Comparing managers and shareholders, it seems that the riskiness of the firm's profit and equity values can be absorbed at lower cost by the shareholders than by the managers. In this view, it appears that the optimal executive compensation plan would place all risk on the shareholders—that is, pay the managers a flat salary. But what about effort?

Trading Off Risk Sharing and Efficiency

We now have two methods owners can use to pay managers. One method aligns the interests of the principal and the agent; that is, contracts are incentive-compatible. The other method assigns risk to the party who can bear it most easily. Can owners integrate these two methods? First they need to account for the fact that firm revenue is risky. Owners can do this by dividing revenue into two parts. The first part, $R_1(e)$, depends on the efforts of managers: The higher the effort of managers, the higher is this component of revenue. The second part is beyond the control of

PROBLEM SOLVED: Two Compensation Schemes Based Solely on Risk Sharing

Consider a firm with risky equity as shown here (note that these values are before deducting for managerial compensation)

Equity = \$10,000,000 (probability 0.5) or
\$20,000,000 (probability 0.5)

Managers are risk-averse. We represent the manager's attitude toward risk by a utility function as shown in Chapter 14. To show risk aversion, we need to show a utility function such as that in panel A of Figure 14.3, in which utility increases at a decreasing rate as wealth increases. This type of concave utility function can be represented by a square root function as follows

$$\text{Manager's utility} = (\text{Wealth})^{0.5}$$

The manager's only wealth is derived from employment. To be competitive, the firm must offer the manager a compensation package that has an expected utility of 1,000; otherwise the manager will leave and find other employment. In contrast to the manager, the shareholders are risk-neutral and interested only in the expected value of the equity after paying the manager's compensation.

Flat Salary of \$1,000,000

The manager has an expected utility sufficient to hold him or her in the job of $(1,000,000)^{0.5} = 1,000$. The shareholders have an expected equity stake, after paying the manager, of

$$0.5(\$10,000,000 - \$1,000,000) + 0.5(\$20,000,000 - \$1,000,000) = \$14,000,000$$

Bonus of a Proportion, x , of the Equity

What proportion x of equity before compensation is necessary to offer the manager an expected utility

of 1,000? We can solve this by setting the expected utility from such a bonus equal to 1,000 (the manager will receive either x times 10 million or x times 20 million)

$$\begin{aligned} \text{Expected utility} &= 1,000 \\ 0.5(10,000,000x)^{0.5} + 0.5(20,000,000x)^{0.5} &= 1,000 \\ x^{0.5}[(10,000,000)^{0.5} + (20,000,000)^{0.5}] &= 2,000 \\ x^{0.5} &= \frac{2,000}{7,634.4} \\ x &= 0.06863 \end{aligned}$$

Note that the expected dollar value of the compensation paid to the manager is now 0.06863 times the expected value of equity before compensation

$$\begin{aligned} \text{Expected value of compensation} \\ &= 0.06863(\$15,000,000) = \$1,029,450 \end{aligned}$$

This risky compensation offers the manager the same expected utility as a flat salary of \$1 million, with the manager being paid a risk premium of \$29,450 to compensate for bearing the risk.

The expected value of equity remaining for the shareholders is

$$\begin{aligned} \text{Expected equity to shareholders} \\ &= \$15,000,000 - \$1,029,450 = \$13,970,550 \end{aligned}$$

Comparing the two compensation plans, we see that both offer the manager an expected utility of 1,000, so the manager is indifferent. However, the flat salary offers shareholders an expected residual equity of \$14,000,000, whereas the bonus plan leaves them with an expected equity value of only \$13,970,550. Therefore, looking only at risk sharing, the flat salary is clearly preferred.

the manager and depends on factors such as interest rates, economic movements, and so forth; we call this \tilde{R}_2 . The tilde over the R shows that this component of revenue is risky. Therefore,

$$R(e) = \underbrace{R_1(e)}_{\text{(under manager's control)}} + \underbrace{\tilde{R}_2}_{\text{(outside manager's control)}}$$

PROBLEM SOLVED: Setting Compensation for Managers

A woman inherits the family farm when she is half-way through her MBA program. She plans to take a job in consulting and is not interested in, nor does she have the skills for, operating the farm herself. However, because the farm has been in her family for generations, she does not wish to sell it. She wants to pass it on to her children, so she decides to hire a manager. Her neighbors tell her that the going salary for a good manager is \$50,000, but she is worried that such a salary will not motivate the manager to run the farm to its potential. She estimates a properly motivated manager could generate profits as shown in the table below. The profit shown in the table is profit gross of compensation costs.

In addition to the manager's effort, profit is also sensitive to the grain price. With low prices, profit is low: Profit increases if grain prices move higher. With her investments and career, she is neutral toward the risk associated with farm income, hence she maximizes profit based on expected value. The manager's only source of wealth is the compensation

from the farm. Being undiversified, the manager is not risk-neutral, so he prefers to maximize expected utility. His utility function if he supplies low effort is

$$U = W^{0.5}$$

If he supplies high effort, it is

$$\begin{aligned} U &= W^{0.5} - u(e) \\ &= W^{0.5} - 46.3 \end{aligned}$$

where U is utility and W is wealth. Note that the effort he expends in working hard costs him 46.3 units of satisfaction. This represents the disutility of effort, $u(e)$.

Flat Salary

Consider first how he behaves if paid a flat salary

$$\text{Utility with low effort} = (50,000)^{0.5} = 223.6$$

$$\text{Utility with high effort} = (50,000)^{0.5} - 46.3 = 177.3$$

Naturally, the manager's expected utility is higher if he supplies low effort. Because the manager is not rewarded for high effort, he will not work hard.

Profit and Managerial Effort

	Grain, Low-Price (Probability 0.5)	Grain, High-Price (Probability 0.5)
Low effort	\$ 50,000	\$150,000
High effort	\$100,000	\$200,000

Owners need to change the compensation structure of managers. Because they cannot observe the effort, nor can it be inferred from profit, compensation cannot depend directly on effort. Owners need to restate compensation as combining a fixed element, K , which is independent of effort, and a bonus, $\alpha\pi(e)$, which depends on profit (and effort):

$$S = K + \alpha\pi(e)$$

Profit-Related Compensation

When managers are paid a straight wage, it is human nature not to work as hard as possible. Why would a manager expend effort if he is not paid to do so? To persuade a manager to work hard, you want to pay him a percentage of profit. If the local labor market for farm managers is efficient, you must offer compensation that generates a net 223.6 units of satisfaction. This would require a flat salary of \$50,000. What percentage x of profit would achieve this? To be competitive, x should satisfy this equation

$$\left\{ \begin{array}{l} \text{Expected utility with} \\ \text{flat salary and low} \\ \text{effort} \end{array} \right\} = \left\{ \begin{array}{l} \text{Expected utility with} \\ \text{bonus of } x \text{ profit and} \\ \text{high effort} \end{array} \right\}$$

$$223.6 = [0.5](100,000x)^{0.5} + [0.5](200,000x)^{0.5} - 46.3$$

$$2(223.6 + 46.3) = x^{0.5}[(100,000)^{0.5} + (200,000)^{0.5}]$$

$$\frac{539.8}{763.4} = x^{0.5}$$

$$x = 0.5$$

Note that the expected compensation would be $0.5[0.5(100,000) + 0.5(200,000)] = 75,000$. The extra \$25,000 above the flat salary of \$50,000 compensates the manager for both disutility of effort and risk.

So if the manager is paid 50% of profit and works hard, he will be just as happy as with a flat salary of \$50,000 and exerting low effort. We still do not know whether the manager will choose to work hard when he gets 50% of profit. To see whether he will, we need to check whether the 50% bonus gives the manager

higher expected utility with high rather than with low effort

$$EU \text{ with high effort} = [0.5][0.5(100,000)]^{0.5} + [0.5][0.5(200,000)]^{0.5} - 46.3$$

$$= 223.6$$

$$EU \text{ with low effort} = [0.5][0.5(50,000)]^{0.5} + [0.5][0.5(150,000)]^{0.5}$$

$$= 216$$

Therefore, the manager will choose to work hard.

The final question that needs answering is whether the owner is better off paying the manager a flat \$50,000 or 50% of profit. This is not a trivial question. If you pay the bonus, he will work harder, and that will increase revenues. On the other hand, you will pay him more on average if you give the 50% bonus (recall that average earnings with this bonus are \$75,000.) To increase net profit, the extra expected revenue from high effort must exceed the additional expected compensation.

The expected net profit will be as follows

$$\text{Flat salary: } E(\text{profit}) = 0.5(50,000) + 0.5(150,000) - 50,000$$

$$= 50,000$$

$$\text{Earnings related: } E(\text{profit}) = 0.5(100,000) + 0.5(200,000) - 75,000$$

$$= 75,000$$

Therefore, the new 50% of profit plan works for both the principal and for the agent.

Correspondingly, here is the profit (note that we deduct only the direct compensation from profit—the remaining compensation, the bonus, is a share of profit)

$$\begin{aligned}\pi(e) &= R(e) - K - C \\ &= R_1(e) + \tilde{R}_2 - K - C\end{aligned}$$

The net benefit of employment for the manager also changes because the manager is exposed to risk if she receives a bonus. The net benefit now comprises two parts. The first shows the expected utility of the manager from wealth. We show wealth as comprising only the manager's employment compensation, S .² The second element is the disutility of effort, $u(e)$

$$\begin{aligned}B(e) &= EU(S) - u(e) \\ &= EU[K + \alpha\pi(e)] - u(e)\end{aligned}$$

2. It is easy to add in other components of wealth such as savings, financial assets, real assets, and so forth.

STRATEGY SESSION: The Value of CEO Pay after the 2008 Crash

The *Bloomberg Businessweek* article leads with the statement: "On many fronts, 2009 is shaping up to be one of the worst years for aligning pay with performance in Corporate America." What's going on? Our pay structure is supposed to be "incentive compatible." What's good for the goose is good for the gander. Performance brings rewards for CEOs because they have performed, but also because they have added value to their stakeholder shareholders.

But several years after Wall Street brought the world to the precipice, many leaders on Wall Street are still receiving 30% more in salaries and bonuses than they did in 2008. Michael Jefferies, CEO of Abercrombie & Fitch, is taking home more pay even though his business performed so badly that the stock plummeted and the firm laid off employees. Because of the appearance of CEOs gone wild, the United States has appointed a "pay czar" who will rule on the compensation packages of seven firms, including Citicorp and General Motors, that received a significant amount of federal bailout money. The government considered establishing controls on the pay packages of all firms that deal with substantial (and perhaps too much) risk

taking—that is, financial firms. Congress also made noise about getting into the act.

Feeling that the best defense is often an offense, some firms are taking proactive steps by cutting compensation at the top. David Chun, CEO of the compensation firm Equilar, claims that "CEO compensation has fallen significantly." Another indicator of these proactive changes, in 2006, 18% of the nation's biggest firms had "clawback" provisions that require top executives to return part of their pay under certain prescribed conditions, such as malfeasance. In 2009, that number rose to 64%.

But let's put the cutbacks into perspective. After the dot.com crash in 2001, CEO pay dropped 14%. After the Great Recession, CEO pay dropped by 8%—this after the greatest U.S. meltdown since the Great Depression.

So how has compensation constituency (base salary, bonus, equity) changed since 2008? For one thing, base salary is not rising, but performance-based bonuses are. In addition, the passage of Sarbanes-Oxley has seen boards of directors become more independent and exercise their independence

The principal–agent incentive conflict can be presented as follows:

1. *The principal's view:* Owners must construct a compensation scheme for managers. They know they cannot perfectly observe managerial actions, so compensation cannot depend solely on effort. However, they can expect the agent to make the effort after she sees the contract terms. Therefore, they want to fashion the compensation to align their interests with that of the agent.
2. *The agent's view:* Having seen the contract terms (particularly how she will be paid), the agent selects a level of effort.
3. *Result:* At the end of the contract, the profit from the agent's activities is revealed. The agent's bonus is paid, and the remaining profit is paid to the principal.

in overseeing executive compensation. For instance, golden parachutes still exist, but the chute only opens two-thirds as wide as before; that is, executives receive two years of salary instead of three. Perks like corporate jets have also disappeared.

Bonuses paid in 2009 for 2008 performance were down 10%. The values of stock options plummeted 15%–25% between 2008 and 2009 because of the poor performance of most companies' stocks. Base salaries were basically flat, as observed by Donald Delves, president of the Delves Group, an executive pay consulting firm.

So what did happen in 2010? As we saw above, compensation fell in 2008 and 2009. But by 2010 it seemed as though nothing had ever happened (if you were a CEO, that is). According to Equilar, the median pay package for the CEO of 334 companies in the Standard & Poor's 500 was \$9 million, which was 25% better than in 2009 (median \$7.2 million) and better than they earned in boomtime 2007 (\$8.4 million). The biggest constituency gainer was in cash bonuses. These bonuses occurred because corporate profits for the 334 firms rose by 41% (mostly via cost cutting and not

by demand increases). In addition, stocks rose 13% in 2010 and have almost doubled in the two years since March 2009. In addition to increasing the value of stock options, the rising stock prices trigger bonuses for some CEOs. The Equilar compensation study looked at companies that had the same CEO in 2009 and 2010.

Who were the big winners in 2010? The highest paid was Philippe Dauman of Viacom, who garnered \$84.5 million (2.5 times his 2009 total compensation—which would have been a mere \$33.8 million). Six of the ten best compensated come from the media and entertainment industries. In second place was Les Moonves of CBS, who was compensated \$56.9 million. In total, the top 10 made \$400 million, up from the approximately \$300 million made by the top 10 in 2009.

Sources: "Executive Pay Trends for 2010," *Bloomberg Businessweek*, October 13, 2009, at www.businessweek.com/managing/content/oct2009/ca20091013-224531.htm; Associated Press, "CEO Pay Exceeds Pre-Recession Level," *CNBC.com*, May 6, 2011, at <http://cnbc.com/id/42929318/>.

STRATEGY SESSION: Principal-Agent Issues in Failed IPOs. Are CEOs Unduly Likely to “Pull the Plug” on IPOs?

An Initial Public Offering, IPO, is an important milestone for many firms. This is their opportunity to access capital markets which offer significant advantages in funding future growth opportunities. But not all IPOs succeed. The most common reason given for failure is poor investor response to the offering, and thus a soft demand for the new share issue. When IPOs fail, there can be significant costs to the firm. Not only does the firm forgo the capital proceeds that had been deemed necessary for growth, but the failure itself provides a negative signal to stakeholders on perceived prospects for the firm.

But sometimes poor investor response is simply due to a generally deteriorating capital market—and it is these external forces that cause the firm to withdraw from its IPO. A recent study by Scott Latham and Michael Braun suggests that cancellation of IPOs might be unduly caused by self-interested CEOs who fail to act in the best interests of their shareholders. This study shows how subtle agency costs can be.

Consider two different forms of the principal-agent problem which are sometimes called “Agency Theory I” and “Agency Theory II.” Agency Theory I holds that the divergence of interests between managers and shareholders will be mitigated as managers hold more equity. Holding equity aligns the

interests of owners and managers and so leads self-interested managers to do what is best for the shareholders. How would this play out in an IPO when capital market conditions are deteriorating? Despite the deteriorating conditions, going ahead with the IPO might still be in the interests of the shareholders. Even if they pay a high price for their new capital, it may benefit them in the long run as the firm can now focus on long-term strategic goals and investments. CEOs with high equity participation have their interests aligned with the shareholders and would likely go ahead with the IPO if they deemed it to be in the shareholders’ interest. However, a CEO without high equity participation might be tempted to abandon the IPO; this way (s)he is able to preserve his/her reputation in the managerial labor market, whereas going ahead with the IPO, and risking its failure, could seriously lower his/her marketability. Such low equity managers would tend to be risk averse.

But the story gets a little more complex with Agency Theory II. This aspect of agency theory picks up another effect of managerial compensation. Managers with high equity participation are typically less diversified than their shareholders. Whereas shareholders often have many assets in their portfolios (indeed many are pension plans or other institutional

The overall situation is illustrated in Figure 15.6. The owners pay the manager a flat salary to provide some certain income to the risk-averse manager. In addition, the manager is given a share of the profit to ensure that she will increase her effort, which will increase expected revenue and profit. Note that no payment relates directly to effort because this is not observable by the firm’s owner. Instead the bonus motivates the manager because higher effort yields higher expected profit, and the manager gets a share of this profit.

investors who hold, on behalf of their members, a large number of assets). In contrast, the typical CEO may hold a large portion of his/her wealth in the equity of his/her employer. A fall in the value of this equity could seriously hurt his/her net worth. A CEO who is risk-averse, and holds a large portion of employer equity, would be unwilling to take the risk of proceeding with an IPO in a down market; this is the same as taking that risk with his/her own net worth. In contrast, under Agency Theory II, the CEO without equity participation has little or no wealth directly tied up in the outcome of the IPO and will not be deterred by the risk of going ahead.

How do these two theories play out in IPOs? Latham and Braun suggest that Agency Theory I implies a negative relationship between the amount of ownership and the tendency for firms to abandon IPOs. Thus, one should expect to find firms where CEOs have low equity participation to frequently abandon IPOs when market conditions deteriorate. But as the ownership increases further the CEO is exposed to a greater and greater undiversified risk. Thus, looking across firms, one would expect that at higher levels of equity participation, the CEO's risk aversion would start kicking in again, thus moderating (and eventually reversing) the negative relation-

ship between equity participation and abandonment of IPOs.

So much for the theory—what happens in practice? Latham and Braun looked at 124 firms that filed to go public at the end of the last tech boom. They then looked at which firms abandoned their IPOs and which ones went ahead. Specifically, they examined the relationship between equity participation and abandonment. And, lo and behold, they found a clear “U” shaped relationship as predicted in their theory.

This study reveals that, for important strategic firm decisions, managerial self-interest may creep in in subtle ways. And while there are partial solutions to principal-agent problems, notably alignment of interests through equity participation, the problems rarely disappear entirely.

Source: Scott Latham and Michael Braun, 2010, “To IPO or Not to IPO: Uncertainty and the Decision to Go Public,” *British Journal of Management* Vol. 21, 666–683.

SOME REFINEMENTS TO MANAGERIAL COMPENSATION

Figure 15.7 shows how managers behave when receiving higher utility with low effort. The higher utility function (red) shows the manager's utility of wealth when supplying low effort. The lower utility function (blue) represents the utility when managers exert high effort. It is shifted downward to show how the disutility of effort lowers overall satisfaction. Thus with a flat salary of B , the manager chooses low effort and realizes utility equals to $U(B)$.

FIGURE 15.6

The Principal-Agent Problem When Effort Is Not Observable

The owner pays a bonus based on profit and flat pay. Because the expected value of the bonus increases with expected effort, the agent increases that effort, which in turn increases the revenue. If the increase in revenue exceeds the increase in pay, then profit increases.

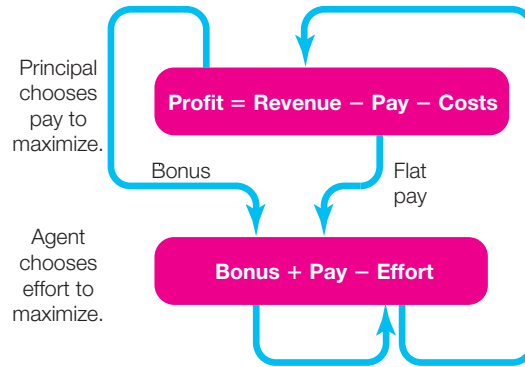
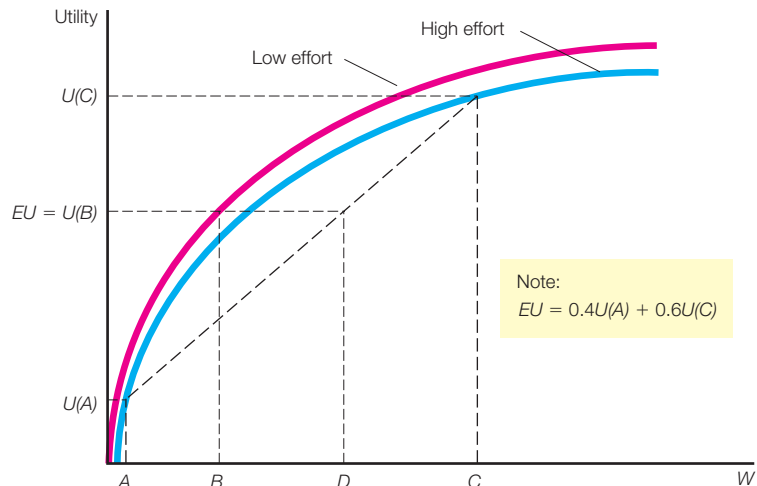


FIGURE 15.7

The Effect of Compensation Schemes on Managerial Effort

The profit sharing gives the manager a 40–60 chance of pay levels A and C , which has an expected utility of EU when the manager supplies high effort. This profit-sharing plan is designed so that working hard gives the manager the same expected utility as a flat salary of B while supplying low effort.



Motivating the Manager with Profit Sharing

Figure 15.7 shows how profit-sharing schemes can affect managerial effort. These schemes give managers a straight percentage of the firm's profit. In this case the manager is paid 2% (1/50th) of profits. The firm's profit is

50 times A with probability 0.4
50 times C with probability 0.6

Being paid 2% of profit, managerial income is

A with probability 0.4 or
 C with probability 0.6

The manager's expected compensation value is $(0.4)(A) + (0.6)(C) = D$ in the diagram. This scheme induces managers to work hard. The expected utility of compensation (blue line) is $(0.4)U(A) + (0.6)U(C)$, which is labeled EU in the figure. This offers exactly the same expected utility as the flat salary of B and low effort, so the manager is just as well off receiving compensation for working hard as having a flat salary and shirking. However, the expected compensation has increased from a flat value of B to an expected value of D . The difference, $D - B$, is the premium necessary to compensate the manager for the risk and disutility of effort.

Motivating the Manager with an Income Guarantee and Stock Options

Note in Figure 15.7 the manager is risking a 40% chance of a fall in income from B to A in accepting the profit-sharing plan. Can we achieve the same effect while protecting the manager from downside risk? Figure 15.8 shows how owners can use stock options. Suppose owners replace the incentive plan with a salary floor (a guaranteed minimum income) of E and a stock option, which has a small chance of paying a large amount of money, F .

The chances of the stock option paying off are 35%. Now the manager's income is either

E with probability 0.65 or
 F with probability 0.35

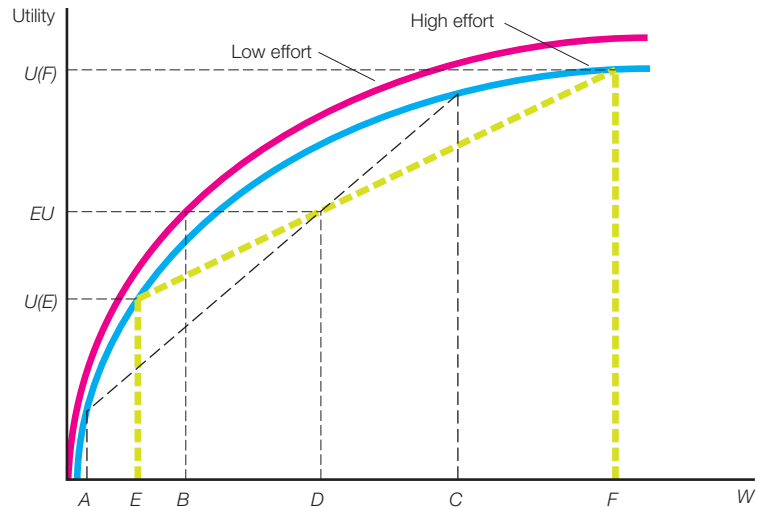
which has an expected value of D (exactly the same as the profit-sharing plan in Figure 15.7). Note that the expected utility from this plan (if the manager works hard) is $0.65U(E) + 0.35U(F)$, which is exactly the same expected utility as from the flat salary of B and shirking.

Why does this scheme reduce the downside risk? First, the manager is paid a flat salary of E . This salary is riskless; it is paid regardless of performance. Second, the manager receives a **call option** on the firm's stock, which is at risk. This option gives managers the right to purchase the firm's shares at some future date. Moreover, the price at which the stock can be purchased (the **strike price** or exercise

Call option Option that gives managers the right to purchase the firm's shares at some future date.

Strike price The fixed price at which the stock can be purchased.

The manager can be provided minimum compensation and strong incentives for efficiency. A flat salary of E is paid together with a stock option, which secures a gain of F minus E if the stock price rises sufficiently. This combination offers the manager the same expected utility as the simple profit-sharing plan in Figure 15.7.



One plan that has attracted attention is the indexed stock option.³ Here the manager is given stock options, but the strike price is not fixed. Rather, it is

592

STRATEGY SESSION: Getting the Board to Focus on the Long Term

In 2006 and (amended in December of 2007), at the urging of Warren Buffett and because of frequent criticism for corporate focus on short-term earnings, the Coca-Cola Company announced a new method of compensating outside Board members (i.e., non-Coca-Cola employees) to the Board of Directors.

Upon appointment to the Board, each outside Board member would be awarded “share units” (basically the number of units of stock that \$175,000 would purchase on February 10, 2006—the average of the high and low stock price on that day); however, these weren’t real stock, they were “share units.” They couldn’t sell them, and they couldn’t convert them into cash unless the firm performed in a particular way in the subsequent three years. If the real stock paid a dividend in the following three years, that dividend was “granted” to the “share units” and would be used to purchase more “share units.” If at the end of three years Coca-Cola’s “amended” value per share had grown by at least a cumulative 8% per year (i.e., by 25.9712% in three years), then, and only then, would the director get paid for his or her service on the 2006 board. If the target is made, board members are paid the cash value of their “share units.” If the target is not made, the board member receives *no* compensation for 2006.

Yet another three-year plan (2007–2009) would determine 2007 compensation and still another (2008–2010) would determine 2008 compensation. After three years, a board member would have three such compensation plans going. The Board of Directors received \$195,617 for 2006 (not paid until February

2009), \$219,861 for 2007 (not paid until February 2010), and \$208,544 for 2008 (not paid until February 2011).

On January 1, 2009, a new compensation plan was adopted and applied to the current Board. The new plan entailed paying the directors a direct cash payment of \$50,000 for the current year’s service and then giving them \$125,000 in “share units” based on the share price on the first business date on or after April 1 (whichever came first). These “share units,” like the earlier ones, allowed for dividends to “purchase” more “share units.”

When a Board member leaves the board, he or she will be paid cash on January 15 of the year following his or her departure, or six months after he or she has left the Board, whichever comes later. Thus, (if Coca-Cola is successful) the majority of a Board member’s compensation is deferred (and for quite some time, for a long-serving Board member). Presumably, the minimum of six months deferral of payment after leaving the Board is to prevent opportunistic departures, such as leaving the Board when the stock value has appeared to peak.

As with the previous plan, the Board believes it aligns the Board’s objectives to those of the shareholders; that is, it maximizes shareholder value (since both are interested in Coca-Cola’s share price) and focuses the Board on the long term—since Board members are not paid (significantly) until they leave the Board.

Source: www.thecoca-colacompany.com/investors/2011_coca-cola-proxy/pdf and the same document for 2006, 2007, 2008, 2009, and 2010.

expressed in relation to an index of stock prices, such as the Dow Jones. As the index rises, so does the strike price of the option. This ensures the manager is not rewarded simply because the market performs well. On the downside, the strike price falls as the market index falls. This ensures the manager is not penalized

STRATEGY SESSION: Call Options

A *call option* is a contract that gives the holder the right (but not an obligation) to purchase a given number of shares of stock in a firm at a preagreed price from a counterparty (the seller of the option). The fixed price is called the *strike price* or *exercise price*.

For example, suppose you have the right to buy x shares in six months' time (known as *maturity*) at a strike price of \$50 per share. Suppose the actual market price of the stock in six months is \$42 per share. You can, if you wish, *exercise your option* to buy the stock (which is now worth \$42) for the strike price of \$50, but that would be silly. You can simply go to your broker and buy the stock at the going market price of \$42. Why use the option to pay more than the stock is worth?

Now suppose the stock rises in price so that in six months it is worth \$74. Your contract allows you to buy the stock for \$50, even though it is now worth \$74. You would *exercise your option* to buy the stock for \$50, thus clearing a profit of \$24 per share.

The value of a call option, and hence its price, depends on several features:

1. The lower the strike price, the higher the value of the option. The holder is paid the difference between the market price at maturity and the strike price (if this difference is positive); thus a lower strike price

increases the payoff and the current value of the call option.

2. The higher the current price of the stock, the higher its value. The higher the current stock price, the more likely the price at maturity will be above the strike price.

3. The higher the risk or volatility of the stock price, the higher the option value. Increased risk enhances the possibility that the price at maturity will be very low (downside risk) or very high (upside risk). Increased downside risk does not hurt the option holder because the option will not be exercised, but increased upside risk enhances the payoff at maturity. Thus, the risk has a one-sided positive effect.

The methodology for pricing options was developed largely by Robert Merton and Myron Scholes, who won the 1997 Nobel Prize in Economics for their work.

In summary, a call option allows managers to make a profit if the stock price rises above the strike price at the maturity date. If the stock price at maturity falls below the strike price, the option is of no value to managers. Hence the option gives managers the chance of gain without the risk of loss. Naturally such options are valuable, and managers must expect to incur a cost (effort) to buy them from the seller.

for poor market performance. The net effect is that the manager is handsomely rewarded if the stock price rises relative to the market. The manager's compensation is now more closely related to factors under his or her control.

PRINCIPAL-AGENT ISSUES IN OTHER CONTEXTS

Similar incentive issues exist between an insured party and the insurance company. These issues are known as **moral hazard**. If you are not insured against fire, car crashes, illness, and other life contingencies, you face the possibility of sudden and possibly crippling financial losses. These uninsured risks should make you

Moral hazard When a party insured against risks behaves differently from the way it would behave if it were uninsured against these risks.

cautious. You should take care in your driving, look after your health, and protect your home by installing smoke detectors and burglar alarms. Although these safety practices can be costly or inconvenient, you are rewarded by lowering the probability, or intensity, of a financial loss. You bear the cost of safety (you pay for the smoke alarm) and reap the reward (you avoid the costs of a fire). If you are insured, there is a separation of the costs and benefits of safety. The policyholder may incur the cost and inconvenience of safety devices and preventive behavior, but the main beneficiary is the insurance firm, which now faces lower expected claim payments.

Insurance moral hazard can be divided into two types: ex-ante moral hazard and ex-post moral hazard. Ex-ante moral hazard refers to the tendency of insured people and firms to take less care to prevent future losses when they have insurance. The absence of smoke alarms relates to ex-ante moral hazard. Ex-post moral hazard is equally important; this refers to the reluctance of policyholders who have already suffered some misfortune to keep the cost of the event under control. Consider a firm that has bought liability insurance against defective products and has now been sued by injured customers. The policy covers both the legal cost of defending the firm against the suit and the cost of compensating the victims. Because the insurer is paying both these costs, some managerial defendants want the settlement to be generous. They see this as a way of buying back customer goodwill at the insurer's expense. Other defendants take the opposite view. Because the insurer is paying legal costs, they want the insurer to spend virtually unlimited amounts to defend the producer's reputation, even if liability seems fairly clear.⁴ Had the managers been without insurance, they might well have sought to balance the incremental costs of defending a claim with the costs of making a settlement offer to the injured customers.

Asset Substitution⁵

Our discussion has centered on the stakeholders in a firm: the shareholders and the managers. But there are other stakeholders under the corporate umbrella, and principal-agent issues also exist between them. We consider now the situation between equity holders and creditors. To focus our attention, we assume the Board of Directors has taken control of managerial compensation and aligned the interests of the shareholders and managers. Therefore, we can be reasonably assured that the managers act on behalf of the firm's owners. The new issue arises because shareholders have gained control (via incentive-compatible compensation) of the decision-making process, but creditors have not.

Consider a drug company with an existing product line that exposes the firm to some risk. Future earnings have an expected present value (*PV*) of either 100 or 200, each with a probability of 0.5. This risk could reflect possible scenarios about consumer demand or the potential for the drug to have unforeseen side effects, resulting in a major lawsuit from injured customers. The firm has borrowed

4. Although almost all insurance policies have limits on what they will pay for damages, some policies carry an unlimited obligation on the part of the insurer for legal defense costs.

5. See Hayne Leland, "Agency Costs, Risk Management, and Capital Structure," *Finance* 53, pp. 1213–1243; and Neil Doherty, *Integrated Risk Management*, chaps. 7 and 8 (New York: McGraw-Hill, 2000).

Because of the financial crisis of 2008, European politicians have declared war on “excessive” compensation. Jean-Claude Juncker (president of the European Commission’s “Eurogroup” of finance ministers) called excessive pay a “social scourge” and demanded that something be done about it. Former French finance minister Christine Lagarde called compensation practices in France “scandalous” and suggested that regulation could be forthcoming. French President Nicolas Sarkozy echoed the sentiment. German president Horst Köhler has denounced high pay.

If it’s excessive in Europe, what would they think about the United States? Let’s look at some work comparing European total CEO compensation with U.S. total CEO compensation.

Is it envy or is it moral outrage? European executives look at their U.S. counterparts and ask: What’s the difference? The usual explanation for why the top 20 highest compensated U.S. executives earn three times the \$12.5 million that the top 20 highest compensated European executives earn has been that the U.S. executives managed bigger companies. But like much conventional wisdom, it isn’t true. The European firms average \$65.5 billion in yearly sales, while the U.S. firms earned \$46.5 billion, so size really doesn’t matter.

Another explanatory piece of conventional wisdom comes from Gordon Gekko (Michael Douglas’ character in the movie *Wall Street*)—greed. During the last 10 years, European executives have felt that U.S. CEO compensation has been excessive. European executives believe their pay is lower because they have concerns regarding social equality. They feel that their American counterparts are greedy.

Pulina Whitaker, leader of the London employment and benefits practice at the King & Spalding law firm, says that her European CEO clients say “that it’s about time that U.S. companies—or shareholders—reined in excessive compensation.” She also sees European and Asian CEOs benchmarking their compensation packages against their U.S. counterparts, but “by and large, they have failed,” notes Whitaker.

In fact, she says that there is a feeling that European CEO pay is now too high (despite the fact that it’s significantly less than that of U.S. CEOs). One reason for this is that in the United Kingdom in the last 10 years, the ratio of CEO compensation to that of the lowest-paid employee of the firm has risen from a multiple of 47 to a multiple of 127 (or 2.7 times). In the United States in 1970, that multiple was 40, but in 2009, it was 344. According to Venkat Venkatasubramanian of Purdue University, in 2009 the European multiple was 20 and in Japan it was 11.

While some argue that it is a world market and hence across the pond(s) benchmarking is relevant, Nicki Demby, a principal in the London office of Towers Watson, the world’s largest compensation consultancy, says that the trend toward worldwide parity with U.S. compensation peaked three years ago (2008) just before the global financial crisis but others (see below) say it’s making a comeback.

In Asia, the golden parachute, a U.S. mainstay, “just goes against the grain of executive mentality in Asia,” according to Miu Hing Alice Au, the head of Heidrick & Struggles’s CEO and board practice in Asia Pacific. “It just doesn’t make sense to CEOs in this part of the world.” In fact, an applicant looking for a golden parachute in Asia would be viewed as unqualified for the CEO position.

Others feel that, with economic recovery, we’re off to the races again with respect to executive compensation worldwide. The basic premise is that it takes a special talent to manage in a global environment. Demand for such managers is high, but supply is low. The law of supply and demand says that compensation will rise when demand exceeds supply at the current price. Korn/Ferry International in Hong Kong reports that it is recommending compensation packages for C-level recruits that are higher than those that Asian companies are used to (and this has been true for the last three years).

The constituent parts of compensation are also changing around the world. Until recently, stock

options were a U.S.-based phenomena. Now, says, Charles Tharp of the Center on Executive Compensation in Washington, D.C., stock-based rewards exist in nearly all international markets (Western Europe and India are cited, with the Middle East coming on board). Ms. Au of Heidrick & Struggles sees the same trend in Asia (except China, so far). One problem is that the equity markets are not well developed yet in some Asian countries; hence, the value of the equity is not easily determined. This is particularly significant because the big difference in CEO compensation between the United States and the rest of the world has been not so much in base salary but rather in deferred stock compensation. An advantage of stock options is the incentive compatibility of aligning the shareholders' and CEO's interests; that is, high performance leads to high stock prices.

John Wilcox, former head of corporate governance of TIAA-CREF, does not believe that we will see parity in CEO compensation across the world. He lists several reasons: Many firms in countries like India have significant family control; some firms have majority control by a group of aligned shareholders; some firms are subject to government control; firms

have different cultures and history (for example, European culture emphasizes equity); firms have different board structures; and some firms allow greater involvement of labor at the board level.

A recent study from the Institute for Compensation Studies at Cornell concludes that, after statistically controlling for some of the variables listed by Wilcox above (for example, firm, ownership, and board characteristics), "U.S. CEOs are paid only modestly more than their European counterparts." In addition, the study finds that pay is more tightly linked to performance in the United States than in most of Europe. And because U.S. executives have been getting stock options for quite some time, U.S. CEOs have more of their wealth in stock and stock options than do their European counterparts. In fact, most of the compensation differential results from stock options (as also concluded above). Favorable tax provisions have encouraged the use of stock options in the United States.

The following table shows the level and constituency of CEO pay in Europe and the United States in 2008 in euros.

As can be seen, the United States is the highest in terms of both average and median CEO total pay,

	Total Pay		Base Salary	Constituency		
	Average	Median		Bonuses	Equity	Other
Belgium	1,328	884	64%	20%	6%	10%
France	1,522	822	60%	21%	15%	4%
Germany	2,606	1,739	39%	42%	9%	11%
Ireland	2,585	1,375	54%	9%	23%	15%
Italy	2,717	2,183	53%	19%	13%	15%
Netherlands	1,526	1,166	49%	21%	17%	13%
Sweden	1,273	1,055	61%	16%	1%	22%
Switzerland	3,636	1,336	57%	17%	12%	14%
United Kingdom	2,016	1,183	46%	18%	28%	9%
European Average	1,989	1,200	50%	21%	19%	10%
United States	3,784	2,414	29%	20%	46%	6%

(Continued)

STRATEGY SESSION: Executive Pay on the Continent (Continued)

just 4% above Switzerland for the average and 10.6% above Italy for the median. Of total pay, the U.S. percentage from base salary is the lowest (at 29%) and 10 percentage points below Germany, which is second lowest. Bonuses were very important in Germany (at 42%), but only half as significant (at 20%) in the United States. Equity is 46% of U.S. CEO total compensation, 18 percentage points ahead of second place United Kingdom. Other compensation is 22% of total compensation in Sweden but only 6% in the United States. When all of Europe is combined as a weighted average of the above countries, U.S. average and median total pay is about double what it is in all of Europe. European CEOs get 50% of their total pay as base salary (whereas the U.S. is only 29%); bonuses are virtually the same between the Europe and the United States, but equity is 46% in the U.S. and only 19% in Europe.

In many European countries, there is now a “say on pay” culture that gives shareholders a say on the pay of top executives. In the United States, rules passed in 2010 require that every three years virtually

every public company must give its investors a vote on what it pays its executives. Although this vote isn’t binding, it is public record, and the power of the press may create pressure. If the shareholders vote no, the company must disclose in its filings to regulators what the resolution was (which might be no change). In 2011, shareholders at only 12 companies have voted against pay plans (this is because large institutional investors tend to side with management). The table below provides some European examples.

Sources: “Pay Attention,” June 12, 2008, at www.economist.com/node/11543665; John Buchanan, “Are CEOs Worldwide Closing the Pay Gap with Their American Counterparts?,” *The Conference Board Review*, Spring 2010, at www.tcbreview.com/dont-look-back/php; Martin J. Conyon, Nuno Fernandes, Miguel A. Ferreira, Pedro Matos, and Kevin J. Murphy, “The Executive Compensation Controversy: A Transatlantic Analysis,” Institute for Compensation Studies, February 13, 2011, at <http://digitalcommons.ilr.cornell.edu/ics/5>; Gerlind Wisskirchen, “Executive Compensation in Europe,” CMS Hasche Sigle, slide presentation, Coronado, California, March, 24–27, 2010.

	Advisory or Binding	What Is Voted On	When Implemented	Regulatory Style
United Kingdom	Advisory, annual	Director’s remuneration	2003	Comply or explain
Netherlands	Binding upon policy change	Binding vote to adopt the remuneration policy for executives and major changes to existing policy	October 2004	Comply or explain
Sweden	Binding, annual	Guidelines for remuneration of senior executives	July 2006	Mixture of comply or explain and regulation
Norway	Binding, annual	Remuneration policy for senior management for coming year	January 2007	Mixture of comply or explain and regulation

Source: Gerlind Wisskirchen, “Executive Compensation in Europe,” CMS Hasche Sigle, slide presentation, Coronado, California, March, 24–27, 2010.

money under a bond issue, and this debt has a **face value** of 100. Even under the worst-case scenario, the firm is worth 100 and can pay back its debt. Therefore, there is no chance the firm will default. We can now see how much the equity and the bonds are worth

Face value The principal amount of the bond.

$$\text{Overall value of the firm} = 0.5(100) + 0.5(200) = 150$$

This value can be divided between the shareholders and the bondholders, with the bondholders' obligations being met first and only what is left (the **residual claim**) going to the shareholders. Therefore, if the firm is worth 100, the debt is paid off and nothing is left for equity. If the firm is worth 200, the debt is paid off, leaving the remaining 100 for the shareholders

Residual claim What is left of the divided value of the firm.

$$\text{Value of bonds} = 0.5(100) + 0.5(100) = 100$$

$$\text{Value of equity} = 0.5(0) + 0.5(100) = 50$$

The firm now faces a new investment decision. It can introduce a new hypertension drug. Its research has come up with two possible formulas. Formula A is moderately effective, has no adverse side effects, and is therefore unlikely to result in any consumer lawsuits. Formula B is a much more effective drug but has greater potential for unwanted side effects. If things go well, the firm could make much more money with formula B. However, if there were a lawsuit, the firm could lose money. The firm must choose which formula to produce. The capital cost of each project is 200 (as shown in Table 15.1), which will be financed by new borrowing. The third column shows the possible value that can be created. With project A, the present value, *PV*, of future earnings are a certain 220. Therefore, the net gain (net present value, or *NPV*) is $220 - 200 = 20$. In contrast, project B could earn a total of only 20 (if there were a lawsuit) or 310 (if there were no lawsuit).⁶ The expected *NPV* is

$$0.5(20) + 0.5(310) - 200 = -35$$

Which project should the firm choose? Clearly project A looks better because its *NPV* is a positive 20, whereas the *NPV* of B is a negative 35. However, the shareholders might look at the decision differently.

TABLE 15.1

Project Selection Using NPV

	Capital Cost	Present Value of Earnings	Expected Net Present Value
Project A	200	220	20
Project B	200	20; probability 0.5 310; probability 0.5	-35

6. The earnings from the projects are independent of those from existing operations.

STRATEGY SESSION: The Good and Bad of Incentive Pay

A number of years ago, Sears auto repair centers offered incentive payments to auto technicians. They were compensated according to the number of repairs they did, such as brake jobs. Sears felt that the technicians weren't productive because they received the same pay whether they worked hard or not. Paying them by the jobs completed would increase their productivity, Sears reasoned. Incentive-compatible, we call it—Sears wants its workers to work hard for Sears' profit, and the technicians want to work hard to increase their compensation. It sounded like a good idea, and normally it is. But in this case, some Sears technicians decided that they could increase their compensation by doing repairs that were unnecessary. An unsuspecting and unknowledgeable consumer who is told that a car's brakes are near failure will undoubtedly authorize a repair. Many consumers were hoodwinked; when the practice was exposed, it caused great embarrassment to Sears.

Do incentive payment schemes (bonuses, stock options) create a similar incentive for potential recipients to cheat? Or do they provide the desired incentive of having them work hard for the objectives of the shareholders?

Some evidence suggests that such compensation packages (especially very large stock options packages) make it more likely that managers will misreport their numbers and make it more likely that stock option grants will lead firms to bankruptcy.

At the Academy of Management meetings in 2005, a paper compared 435 companies that were forced to restate their financial statements with comparable companies that did not have to restate their earnings. The greater the proportion of managerial pay in stock options, the greater was the probability the firm had to restate its profit. In cases where bosses received more than 92% of their compensation in stock options, about 20% misstated their profit in a five-year period. If

the firm isn't doing well, the tendency is to fake results to get compensated and to keep those who employ you happy. Likewise, if you had a good year last year, an even better year will be expected of you this year. The pressure is on.

A study by Moody's (the bond rating service) revealed that firms with the highest-paid bosses (controlling for company size and performance) experienced a greater probability of defaulting on debt or experiencing a large fall in their bond ratings. How do we know whether executives in the Moody's case were cheating? Perhaps the executives were risk takers, and the investments didn't work out. Alternative explanations include weak board oversight and that incentive pay packages "create an environment that ultimately leads to fraud." WorldCom and Enron are examples of the latter.

But do boards of directors offer enough protection against abuse? Not always. A *New York Times* editorial asked, How irresponsible does an outside director have to be before he/she faces legal responsibility for not performing duties? The editorial concluded that the answer is very, very irresponsible. It went on to imply that some boards blithely take their (not insubstantial) fees and ignore their duties. Two examples were cited: A Delaware judge harshly criticized Disney's board for the hiring and firing of Michael Eisner but concluded that no illegal acts occurred. A Krispy Kreme Doughnut's special committee of the board (composed of board members appointed after a major fall at Krispy Kreme) concluded that the prior directors did nothing illegal. Many feel that the Board had watched from the sidelines as the problems of the company occurred and rubber-stamped many decisions that should have required diligent oversight. The special committee stated that the board didn't oversee management's processes and decisions with an appropriately skep-

tical eye—concluding, however, that they did nothing they knew or believed was in clear conflict with the best interests of Krispy Kreme and likewise did nothing where they received improper personal benefit. These are not the only boards that did not practice due diligence and yet bore no legal consequences.

And just how strongly are the performance goals written and adhered to when it comes time to grant compensation? At the Las Vegas Sands Corporation, five top executives were overpaid (relative to the company's contracts with them) by over \$2.8 million in 2005. The board said that it was okay because of the company's outstanding performance (despite the fact the capitalized value of the corporation fell by 18%).

It apparently is common to find instances where compensation exceeds the amounts allowed under the performance contracts approved by boards. Other examples include Halliburton, Assurant, Mothers Work, and Big Lots. Critics are concerned that standards are lowered after being initially set (so that they are easier to attain) and that they are not set high in the first place.

Some compensation packages seem to create perverse incentives. At the News Corporation, the COO was to receive a bonus of \$12.5 million if earnings exceeded 15% in any given year but a bonus of \$4.5 million if earnings *fell* by 6.25% and \$3.52 million if earnings *fell* by 14%. Thus, even if the COO took off a year from managing, he would still be far from poor.

One way to circumvent transparency in bonus calculation is to have a list of vaguely defined measures that will trigger bonus payments. In addition, a board may weight these items in any way it wishes (including a zero weight). The reason for such published ambiguity, it is stated, is that rivals could calculate an executive's compensation, figure out secret company data, and put together a compensation

package to woo away an executive. Critics say that carefully chosen measurement variables can trigger bonuses whether the overall firm does well or not.

An insurance company executive was well compensated despite losses to the company because of the extraordinary events caused by hurricane claims. The hurricane was called an act of God; the same act of God didn't spare the company's shareholders from a loss of capitalized value of their holdings. Owners take risks by owning a company (profits can go up or down), but executives that work in the insurance industry must take the risk that acts of God occur and may impact their compensation. When the hurricane impact was stripped out, the insurance company paid bonuses 1.72 times their target.

Other executives were paid bonuses for work that would seem to be part of their general duties, such as compliance with the Sarbanes-Oxley Act. At Mothers Work and Big Lots, bonuses were paid although both firms did not achieve the targets that would trigger the bonuses. They were paid for reasons such as other activities achieved but not listed in the performance criteria for bonus payment, to get management payroll up to the industry standard, and to keep the management team intact.

Who watches the watcher who watches the watched?

Sources: Floyd Norris, "Stock Options: Do They Make Bosses Cheat?" *New York Times*, August 5, 2005, at www.nytimes.com/2005/08/05/business/05norris.html; Floyd Norris, "Inept Boards Need Have No Fear," *New York Times*, August 12, 2005, at www.nytimes.com/2005/08/12/business/12norris.html; Gretchen Morgenson, "Big Bonuses Still Flow, Even If Bosses Miss Goals," *New York Times*, June 1, 2006, at www.nytimes.com/2006/06/01/business/01bonus.html.

Value of the Firm If Project A Is Chosen

Consider the total value of the firm with each project choice. With project A, the firm has either 100 or 200 from existing operations plus an additional 220 from the project. This gives a total value of either 320 or 420. The total value is divided between the bondholders, to whom the firm owes a total of 300 (the original debt was 100, plus the firm borrowed 200 for the new project), and the shareholders. We assume that the original 100 must be paid first because it was borrowed first (it is called senior debt). The debt raised to fund the new project is junior and can be paid only after the original debt is paid.

Note in Table 15.2 that the minimum value of the firm is 320, but it owes a total of 300. Therefore, there is always enough value to pay off the debt in full.

Value of the Firm If Project B Is Chosen

If project B is chosen, the total value can be either 100 or 200 from existing operations plus either 20 or 310 from the project. This leaves possible total values of

$$\begin{aligned} 100 + 20 &= 120 \\ 100 + 310 &= 410 \\ 200 + 20 &= 220 \\ 200 + 310 &= 510 \end{aligned}$$

The values of the firm and of debt and equity are as shown in Table 15.3.

Here is the problem. The firm borrows 200 to spend on one of these two projects. Then shareholders must choose which project to undertake. The choice should be clear. Shareholders are better off with project B, where their equity is worth 80, than with A, where their equity is worth only 70. How can what seems like the worse project actually make shareholders better off? The problem has to do with limited liability. If the firm undertakes A, it creates no risk of defaulting on the debt. Shareholders gain the full *NPV* of the new project of 20 (note that before the project is undertaken, equity is worth 50; after project A, equity is worth 70). But if B is undertaken, there is a 50% chance that the project will fail, resulting in

TABLE 15.2

Firm Value If Project A Is Chosen

Value of the firm	$0.5(320 + 420)$	$= 370$
Old debt	$0.5(100 + 100)$	$= 100$
New debt	$0.5(200 + 200)$	$= 200$
Equity	$0.5(20 + 120)$	$= 70$

TABLE 15.3

Firm Value If Project B Is Chosen

Value of the firm	$0.25[120 + 220 + 410 + 510]$	= 315
Old debt	$0.25(100 + 100 + 100 + 100)$	= 100
New debt	$0.25(20 + 120 + 200 + 200)$	= 135
Equity	$0.25(0 + 0 + 110 + 210)$	= 80

a value of only 20. If that happens, the firm is bankrupt and unable to fully pay off the debt. On the other hand, if the project succeeds (is worth 310), the shareholders reap a big reward. Therefore, when things go well, the shareholders keep all the upside risk; but when things go badly, the shareholders walk away from the debt. Shareholders are playing a “heads we win, tails bondholders lose” strategy.

This illustrates an important principal-agent issue. When firms have a significant amount of debt, the shareholders tend to favor unusually risky investment decisions. It seems that the bondholders are the unwilling victims of these games. But let us look at the bondholders for a moment; they are not totally helpless. Bondholders must decide whether to lend to the firm and how much they wish to pay for the bonds. In this case the firm is trying to issue bonds with a face value of 200. If you were an investor looking at this firm, your thoughts might progress as follows:

If I were to pay 200 to buy these bonds, what would the shareholders choose to do with the money? Well, the rational thing for the shareholders to do once they have my money is to choose project B because the shares would then be worth an expected price of 80 (compared with 70 for project A). In that case I should anticipate that B would be chosen, and my new bonds would be very risky and worth only 135 (see the value of “new debt” in Table 15.3 when B is chosen). Consequently I would be unwilling to buy these bonds for 200; rather I would pay only 135 for them, which is what they are worth.

If we follow the logic a little further, we see that managers are unable to undertake either project. Because rational bondholders anticipate that B would be chosen, they are willing to pay only 135 for the new debt issue, even though the face value is 200. Because the capital cost of project B is 200, the amount raised from the debt issue is insufficient to fund the project. Therefore the firm is unable to finance project B. Does that mean that A would be chosen? Suppose, indeed, that the managers announced their intention to choose project A. Unfortunately

STRATEGY SESSION: The Song of the Sirens

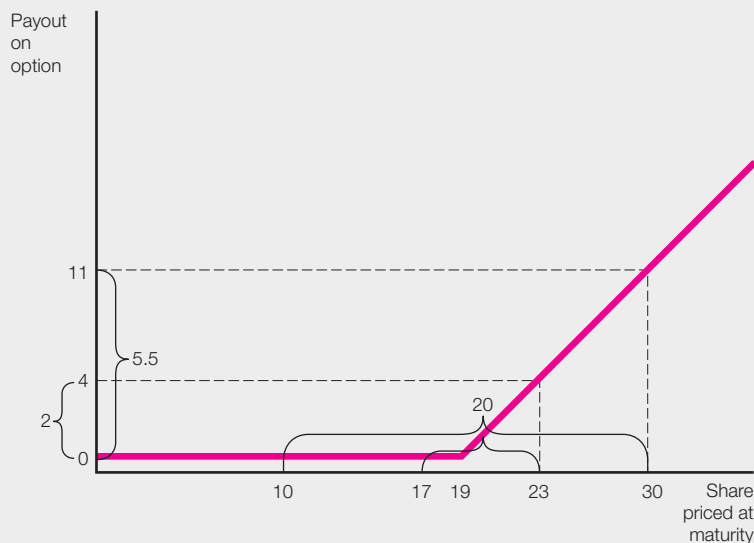
A defining feature of the 2008 financial crisis was the implosion of financial institutions due to what looks like excessive risk taking. It was a quick end for some (Bear Sterns, Lehmann, Wachovia, Northern Rock Plc) and major restructuring and government aid for others (AIG, Citicorp, Fortis).

Managers of these institutions were taking increasingly complex and risky securities on their books. Indeed the level of risk of these securities can only be estimated with very sophisticated mathematical modeling and even then, the models may underestimate the true level of risk.^a Often, managers of these institutions magnified the inherent risk in the securities they held by using a financial tactic called “leverage,” whereby they borrow money to increase their holdings. Some investment banks were leveraged at a 30-to-1 ratio. For every one dollar of equity, managers held \$30 of debt. In using debt, managers take on the additional risk of losing control of assets, because they are unable to repay the principal or

interest on that debt. By historical measures, the leverage risks managers at many institutions were taking on were of unusually high levels.

Many observers believe compensation schemes encouraged managers to take on additional risk. This, in turn, caused them to engage in moral-hazard-like behavior. Consider the role of stock options, a common component of managerial compensation. A *stock option* is the right to purchase the firm’s shares at some time in the future at an agreed price. The executive might receive a number of these options as part of his overall compensation package. We show how these are valued in the Strategy Session titled “Call Options” on page 594. In theory, these options should increase managerial focus on long-term results since they expire at a future date.

From a behavioral point of view, options give managers all the upside potential from rising share prices but none of the downside loss if the share price should fall. Consequently, *these options are more valu-*



able the greater the riskiness of the underlying stock. This is illustrated in the previous diagram. Consider a firm whose share price in one year could be either 17 or 23 each with a 50% probability. The expected share price is therefore $0.5(17) + 0.5(23) = 20$. Now an option with a strike price of 19 will have the following payoffs:

- If the share price at maturity is 17 option pays 0 since the option is not exercised.
- If the share price at maturity is 23 option is exercised and has net payoff of $23 - 19 = 4$.

Thus, the expected payout to the manager on the option is $0.5(0) + 0.5(4) = 2$.

Now suppose the manager makes a different set of decisions that result in a very different risk profile. With these alternative decisions, the share price at maturity has a 50% chance of being 10 and a 50% chance of being 30. Notice the expected share price is the same as before, i.e., $0.5(10) + 0.5(30) = 20$. Now an option with the same strike price of 19 will have the following payoffs:

- If the share price at maturity is 10 option pays 0 since the option is not exercised.
- If the share price at maturity is 30 option is exercised and has net payoff of $30 - 19 = 11$.

Thus, the expected payout to the manager on the option^b is $0.5(0) + 0.5(11) = 5.5$. By changing the riskiness of the share price without changing its expected value, the manager has substantially increased the payoffs he might receive from the option when it matures.

And herein lies the source of this moral-hazard-like behavior. Stock options provide managers with powerful incentives to make decisions that increase

the stock price, which is why they are used in the first place. However, the pernicious side effect is that they also provide incentives for managers to increase the riskiness of the stock, for this too increases the value of the option.^c

Stock options were widely used in financial institutions. In fact, most of the firms that received government support paid the highest compensation to their CEOs. Lloyd Blankfein, the CEO of Goldman Sachs received \$54 million in 2007; JPMorgan Chase CEO, James Dimon took home \$30 million that year; while G. Kennedy Thompson, CEO of Wachovia received over \$15 million in 2007.^d If you believe CEOs respond to incentives, it should not be a surprise that managers sought to maximize the value of their options by trading in such risky securities and then leveraged that risk with debt financing.

^aOne of the biggest of all collapses was that of the insurance giant AIG. In addition to its regular insurance business, AIG had a special unit that "insured" mortgage debt through complex financial instruments known as credit default swaps. Despite having a very sophisticated, state-of-the-art, mathematical model to track its exposure, AIG was still unable to avoid a catastrophic failure that required a Federal Government bailout. See "Behind AIG's Fall, Risk Models Failed to Pass Real-World Test," Carrick Mollenkamp, Serena Ng, Liam Plevin & Randall Smith, *The Wall Street Journal*, October 31, 2008.

^bWe need a caution on the price of the option. Pricing is quite complicated and is not simply the expected payout.

^cSee Neil Doherty, *Integrated Risk Management* (McGraw-Hill, 2000) for a discussion of this incentive mechanism. There are many empirical studies confirming that executives who are paid with stock options do tend to make riskier decisions. See for example Peter Tufano, "Who Manages Risk: An Empirical Examination of Risk Management Practices in the Gold Mining Industry," *Journal of Finance*, Vol. 51, pp. 1097-1137.

^dCecilia Kang and Annys Shin, "Top Executives at Bruised Firms among Wall Street's Highest Paid," *The Washington Post*, September 24, 2008, p. D03.

investors buying the new debt issue would still rationally assume that if they subscribed 200 for the issue, the shareholders would change their minds and use the 200 to fund project B. Hence investors still would subscribe only 135 for the new bond issue. The firm is unable to undertake either project. Because bondholders anticipate the shareholders' temptation to choose the risky negative *NPV* project, the firm is unable to fund either project. It is forced to sacrifice not only the expropriatory project B but also a project with a genuine positive *NPV*.

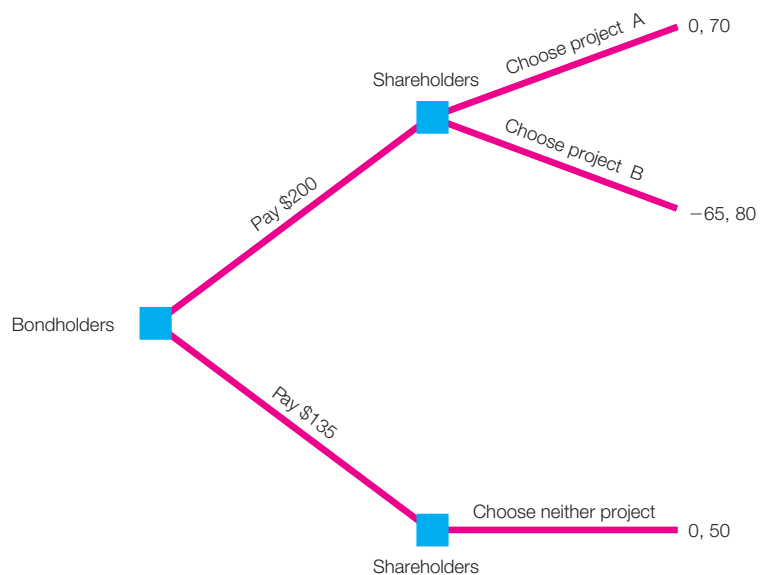
Representing the Bait and Switch in Game Theory

Figure 15.9 shows this issue in a game theory form. The shareholders must choose project A or B. Bondholders must choose either to pay the full 200 for the new debt or only 135. Because managers cannot undertake either project unless bondholders pay the full 200, shareholders have only the original equity value of 50 (from the existing product line). However, if the bondholders pay the full 200 to cover the project costs, the shareholders have equity value of 70 from choosing A or 80 from choosing B. The bondholders' payoff is the difference between what they pay

FIGURE 15.9

Will Shareholders Pull the Bait and Switch?

If bondholders pay 200, they anticipate that shareholders will choose project B. These bondholders will pay only \$135 for the debt. Note that the first number at the end of each branch is the payoff for the bondholder and the second number is the payoff for the shareholder.



for the bonds when they are issued and what they are worth. If the bondholders pay 200 and the firm chooses A, the bonds are worth 200 and the net payoff is 0. If the bondholders pay 200 and the firm chooses B, the bonds are only worth 135, leaving a net payoff of -65 . If the bondholders anticipate the bait and switch, they pay only what the bonds are worth and have a net payoff of zero.

The game is sequential, with the bondholders making the first move. The bondholders anticipate that the shareholders will choose B if they pay 200. Therefore the bondholders pay only 135, and the shareholders cannot undertake either project and are left with equity of 50. Shareholders would like to make the following promise to the bondholders: If you pay 200, we promise to undertake project A. The problem is that the promise is not credible. Having received the 200, the shareholders would then have an incentive to change their minds. There is a credibility problem; shareholders have an incentive to pull a bait and switch.

Possible Solutions to the Asset Substitution Problem

There are several ways managers can avoid or minimize this problem:

- *Fund with equity*: The problem arose because managers tried to use debt to fund the project. If they can pay for the project by using internal funds or by making a new issue of shares, then the problem is mitigated or disappears.
- *Establish a reputation for protecting creditors*: If the firm follows a consistent policy of making decisions that preserve the interests of creditors, then a promise not to undertake risky projects in the future might be viewed as credible.
- *Precommit to hedge or insure risk*: Another way this problem is sometimes tackled is that the firm voluntarily commits itself to insure the risks. The problem arose because project B was risky, and this risk caused the possibility of default on the debt. If the risk were insured, then bondholders would be protected. Many debt instruments carry a legal obligation for the firm to insure assets.

PRODUCT LIABILITY AND THE SAFETY OF CONSUMER GOODS

Many countries have laws that protect consumers against the risks of injury from defective products. These laws serve two purposes. First, they compensate injured consumers. If a consumer is injured by a defective product, the producing firm has to pay compensation. Second, because the firm must pay compensation if products are defective, there is an incentive for the firm to make safer products. Safer products result in fewer costly lawsuits. These laws solve a principal–agent issue. The firm makes decisions on safety, but the consumer bears the costs if the product causes injury. The product liability law is incentive-compatible because it aligns the interests of the principal (the consumer) and the agent (the producer).

STRATEGY SESSION: Moral Hazard and the National Health Care Debate

Much of the national debate over health care reform has centered on moral hazard and principal-agent issues. The moral hazard issues in health care are many. For example:

- Is a doctor's choice of treatment of a patient influenced by the way he/she is paid? For example, will doctors paid on a "fee for service" basis tend to prescribe more treatments than one paid on "capitation" (a fixed sum per patient that does not vary with the level of treatment).
- Are physicians who have an ownership interest in expensive equipment (e.g. MRIs) more likely to prescribe their use than doctors who have no such interest?
- Are patients who have more insurance—lower copays—prone to make more doctor visits than those with less insurance?

On this last issue, consider a recent article in the rather liberal online magazine, *Slate*. The article notes that co-pays in health insurance have increased enormously in recent years and, as economic theory would predict, this has led people to ration their access to health care. The important issue is whether this rationing is a reduction in the number of frivolous visits to the doctors, which would be a good thing, or whether co-pays discourage those who are really ill from seeking treatment (a bad thing). Liberals and conservatives tend to differ in their emphasis; conservatives stress the former and liberals the latter.

The author notes that the elderly may be particularly prone to frequent doctor visits, if only because of their high volume of minor ailments. But, the article quotes from a recent New England Journal of Medicine study of elderly patients:

Assuming an average reimbursement of \$60 for an outpatient visit, seven annual outpatient visits per enrollee, and an average copayment increase of \$8.50 per visit, a Medicare plan would receive an additional \$5,950 in patient copayments and avert \$1,200 in spending on outpatient visits for every 100 enrollees, for a total of \$7,150 in savings for the health plan.

This is where discussions of moral hazard typically begin and end.

However, assuming an average cost of \$11,065 for hospitalization of a person 65 to 84 years of age in 2006, our estimates suggest that expenditures for inpatient care will increase by \$24,000 for every 100 health plan enrollees in the year after copayments for ambulatory care are increased. Save \$7,150, spend \$24,000. Net loss: \$16,850.

Sources: Timothy Noah, "Health Reform and Moral Hazard: Would Health Reform Boost Frivolous Doctor Visits?" *Slate*, Feb. 3, 2010; Amal N. Trivedi, M.D., M.P.H., Husein Moloo, M.P.H., and Vincent Mor, Ph.D., "Increased Ambulatory Care Copayments and Hospitalizations among the Elderly," *New England Journal of Medicine*, 2010; 362:320–328

How Safe Would Products Be without a Product Liability Law?

How much should managers invest in making safer products. We use s to denote the expenditure undertaken to make safer products. Therefore, the cost of safety is s , and the marginal cost is simply 1

Total cost of safety = s

Marginal cost of safety = $\frac{\Delta s}{\Delta s} = 1$

The benefit of safety is that it reduces the expected cost of injuries to consumers. Assume the expected cost of such injuries is

$$\text{Expected cost of accidents} = 4,000 - 20s^{0.5}$$

We can see that the expected cost of accidents depends on the firm's choice of s . Moreover, as s increases, the expected cost falls. This is represented as the marginal benefit of safety; because the benefit of safety is a reduction in a cost, we must remember to change the sign

$$\text{Expected marginal benefit of safety} = \frac{\Delta[-(4,000 - 20s^{0.5})]}{\Delta s} = \frac{10}{s^{0.5}}$$

QUANT OPTION

Because the cost of safety is s , the marginal cost of safety is $ds/ds = 1$. The expected benefits of safety are just the negative of the expected cost of accidents—that is, $Z = -4,000 + 20s^{0.5}$. The expected marginal benefits of safety are then $dZ/ds = 10/s^{0.5}$.

However, in the absence of a product liability law, the firm is not required to compensate consumers for injuries suffered, and these costs fall directly on customers. The expected benefit to managers for safer products is zero. Hence the firm pays all the costs of safety and receives none of the benefits. The profit-maximizing firm chooses $s = 0$. In this case products are unusually dangerous.

Safety under a Product Liability Law

The product liability law can be rationalized as providing an incentive for managers to make safer products. To achieve this, the costs of accidents are borne by the firm, which now must compensate the victims. Therefore, the expected benefit to the firm from spending s on safety is the reduction in the expected accident cost. The level of safety now optimal for the firm is that which equates marginal cost and marginal benefit

$$\text{Marginal cost} = \text{Marginal benefit}$$

$$1 = \frac{10}{s^{0.5}}$$

$$s = 100$$

The introduction of the law increases the firm's choice of safety from 0 to 100. Consequently the expected cost of accidents declines from $[4,000 - 20(0)^{0.5}] = 4,000$ to $[4,000 - 20(100)^{0.5}] = 3,800$. This obviously looks like a social benefit; but was the law really necessary to achieve this?

Optimal Safety under a Market Mechanism

Clearly the product liability law is incentive-compatible. But would an incentive-compatible solution evolve under private market incentives? There is a marketplace for information. Specifically, information about product safety is widely disseminated in subscription magazines, newspaper articles, and television news shows. The volume of such information is increasing, and cost and access to information are improving as it becomes available over the Internet. Moreover, if there were no product liability law, such information would be more valuable and perhaps even more widely available. Without a product liability law, consumers would bear the costs of accidents themselves, and safety would be reflected in the demand for, and therefore the price of, products.

Suppose managers sell a product in a quantity of 1,000, and customers are willing to pay a price of 30 if they know the product is perfectly safe. The firm's revenue would then be $30 \times 1,000 = 30,000$. The total cost of the accidents is $[4,000 - 20(s)^{0.5}]$. With 1,000 customers, the cost to each customer is $[4,000 - 20(s)^{0.5}]/1,000 = [4 - 0.02(s)^{0.5}]$. Therefore, while the customer will pay 30 if she knows the product is safe, she will discount that price by the expected cost of accidents if information is available about product safety. Customers are willing to pay a price of only

$$\text{Price} = 30 - (4 - 0.02s^{0.5})$$

The total revenue becomes

$$\text{Total revenue} = 1,000[30 - (4 - 0.02s^{0.5})] = 30,000 - (4,000 - 20s^{0.5})$$

The revenue is now discounted by exactly the expected cost of accidents. Assume the firm has other costs of production of 10,000. The profit of the firm is now

$$\begin{aligned}\pi &= \text{Total revenue} - \text{Cost of safety} - \text{Other production costs} \\ \pi &= 30,000 - (4,000 - 20s^{0.5}) - s - 10,000\end{aligned}$$

The firm must choose the level of safety that maximizes profit, so we can set $\Delta\pi/\Delta s$ equal to zero

$$\frac{\Delta\pi}{\Delta s} = \frac{10}{s^{0.5}} - 1 = 0$$

When solved for s , this gives an investment in safety of $s = 100$.

Of course this is exactly the same result we got with the product liability law. This should not be a surprise. With product liability, the firm paid the full expected cost of accidents directly in lawsuits. Under the market mechanism, the firm had its revenue reduced by the full expected cost of accidents. Either way, the firm internalized the full cost of accidents and had an incentive to reduce these costs by investing in safety.

STRATEGY SESSION: Moral Hazard in the Financial Market: The Rescue of the Investment Bank Bear Stearns

In 2007 Bear Stearns's stock had been trading as high as \$170. But the subprime mortgage crisis severely impacted Bear Stearns, which had provided lines of credit to subprime lenders and also was the owner of EMC Mortgage. By mid-March 2008 the share price was down to \$30, and fearing failure of the bank and its impact on the financial system, the Federal Reserve and rival firm JPMorgan Chase arranged a rescue plan (often called a "bailout"). Under this plan, the Fed would provide a limited guarantee of Bear Stearns' obligations, and JPMorgan purchased its rival for \$10 per share. JPMorgan would bear the first part of the Bear Stearns obligations before the federal guarantee kicked in.

This is an interesting moral hazard story. Bear Stearns got into this mess by accepting high-risk activities and by not managing that risk effectively (by appropriate diversification and hedging strategies). Although the Fed perceived that a Bear Stearns failure could create a financial meltdown, bailing out the firm created an expectation that large financial institutions would not be allowed to fail (at least not when financial markets are uncertain). Will such bailouts encourage institutions in the future to take big risks? The financial system relies on prudent financial behavior. If the imprudent are allowed to avoid the dire consequences of their risky actions,

they are hardly likely to exercise prudence going forward. Thus such bailouts create adverse incentives and may increase financial risk taking (and therefore the risk to the whole financial system). Bailouts thereby create a moral hazard.

The counterarguments are also couched in moral hazard terms. By receiving only \$2 per share, Bear Stearns's shareholders were hardly being bailed out—they were indeed taking an enormous loss even considering the share price of just a few days earlier. In fact the offer was increased to \$10 per share to make it more acceptable. This might increase the moral hazard. However, this was offset, to some extent, by reducing the value of the Fed's guarantee. Thus we can see the whole exercise as one of trade-offs. The Fed clearly did not want the bailout to be too attractive and required that the shareholders suffer considerable pain. It wanted to rescue the financial system from its mess—but without creating expectations of generous bailouts in the future.

Source: Andrew Ross Sorkin, "JP Morgan Pays \$2 a Share for Bear Stearns," *New York Times*, March 17, 2008, at www.nytimes.com/2008/03/17/business/17bear.html.

QUANT OPTION

The firm will maximize profit if

$$d\pi/ds = (10/s^{0.5}) - 1 = 0$$

This shows that there are generally different methods to address principal-agent issues. One way is through the force of civil law, making firms responsible for defective products and therefore aligning the interests of firms and their customers. Another way is to appeal to market mechanisms by which self-interested customers seek out information and use this information in their purchasing decisions. Customers reward (or punish) firms by varying the price they are willing to pay according to the level of product safety. Thus price becomes a way of aligning the interests of the firm and its customers.

SUMMARY

1. A principal-agent relationship is one in which a principal employs an agent to undertake a task (such as shareholders employing managers to run a firm). The objectives of the principal and agent may be quite different. Shareholders like a high profit and high share price. Managers like such things as prestige, income, pleasant work, and perquisites. More simply, managers may wish to do as little as possible for the greatest possible reward. Therefore, we use the term effort to describe a manager forgoing his or her own objectives (the manager does not shirk) to attend to the wishes of shareholders. Greater effort on the part of the manager can usually increase a firm's profit.

2. The problem for the principal is to motivate the agent to work for the benefit of the principal. Shareholders seek to get managers to supply great effort and maximize the firm's profit. If the shareholders could observe the effort of managers, this would be no problem. The managers' compensation could be scaled to the managers' effort. However, if the shareholders do not know how hard managers work, they are hard-pressed to reward them directly for effort.

3. The full principal-agent problem arises because the firm's owners lack the time or skill to observe all the actions of the manager. Owners cannot perfectly observe the manager's effort. Nor can they infer accurately the effort from the firm's revenue or profit if that profit is risky. The solution recognizes that, on average, greater effort results in higher profit. Managers are motivated by being given a share of the profit or an equity stake in the firm. However, this incentive compensation is risky for the manager. Profit can vary for reasons outside the manager's control. Therefore, compensation usually has a fixed component and a profit-related portion.

4. A particularly powerful type of incentive compensation is the stock option. The manager is given an option to buy the firm's stock at some future date at a pre-agreed price. These options are risky for the manager. If the stock price falls, the options are worth nothing, but a big increase in the stock price can bring fabulous returns. The manager is penalized severely for bad firm performance

and rewarded handsomely for good performance. In the bull stock market of the 1990s, stock option plans brought great wealth to many managers.

5. Another type of principal–agent problem arises between shareholders and the firm’s creditors. This is known as asset substitution. Limited liability means that when a firm is bankrupt, the shareholders can walk away rather than pay the creditors. This creates an incentive for the shareholders to take on risky investments. Risk implies a chance of a large upside gain and a chance of a large downside loss. If luck is favorable, the creditors can be paid off, and all the upside gain goes to the shareholders. But if things turn out badly, the shareholders can use bankruptcy law to default on the debt. Thus high-risk projects have a “heads I win, tails you lose” quality for the shareholders. Of course risky projects hurt the creditors. This tension may lead the firm into dysfunctional investment decisions, and the firm may be forced to limit its debt financing to resolve this type of agency problem.

6. The provision of an optimal level of product safety can be viewed as a principal–agent problem. Product liability laws require firms to pay for damages caused by their products. The probability of making these payments creates an incentive for firms to make their products safer. The same level of safety can be attained by the marketplace if consumers can be cheaply informed about a product’s safety level. Unsafe products would then sell at a discount relative to safe ones, and firms would be motivated to improve safety to increase their profit.

PROBLEMS

1. Your business generates the following profits (these are stated before compensation is paid to the manager):

	Low Demand (0.3)	Medium Demand (0.4)	High Demand (0.3)
Low effort	\$5 million	\$10 million	\$15 million
High effort	\$7 million	\$12 million	\$17 million

You see that profit depends on both the level of effort chosen by the manager and the level of demand. The demand level is random, and the probabilities of each demand level are shown. So with low effort, expected profit is \$10 million; with high effort, it is \$12 million. The manager has a utility function that is either

$$\begin{aligned} \text{Utility} &= (\text{Wealth})^{0.5} \text{ if effort is low or} \\ \text{Utility} &= (\text{Wealth})^{0.5} - 100 \text{ if effort is high.} \end{aligned}$$

Therefore -100 is the disutility of effort. You are interested in maximizing the expected profit after deduction of compensation. You consider three different compensation packages:

- A flat salary of \$575,000.
- A payment of 6% of profit.
- A flat payment of \$500,000 plus half of any profit in excess of \$15 million.

Which compensation do you choose?

2. Suppose the typical Florida resident has wealth of \$500,000, of which his or her home is worth \$100,000. Unfortunately Florida is infamous for its hurricanes, and it is believed there is a 10% chance of a hurricane that could totally destroy a house (a loss of \$100,000). However, it is possible to retrofit the house with various protective devices (shutters, roof bolts, and so on) for a cost of \$2,000. This reduces the 10% chance of a loss of \$100,000 to a 5% chance of a loss of \$50,000. The homeowner must decide whether to retrofit and thereby reduce the expected loss. The problem for an insurance company is that it does not know whether the retrofit will be installed and therefore cannot quote a premium conditioned on the policyholder choosing this action. Nevertheless, the insurance company offers the following two policies from which the homeowner can choose: (1) The premium for insurance covering total loss is \$12,000 or (2) the premium for insurance covering only 50% of loss is \$1,500. The typical homeowner has a utility function equal to the square root of wealth. Will the homeowner retrofit the house, and which insurance policy will the homeowner buy? Will the insurance company make a profit (on average) given the homeowner's choice?
3. The expected profit of your firm is 1,000, plus 500 if the manager works hard. The manager receives a flat salary of 100 plus a portion x of any profit in excess of 1,300. The manager's utility function is

$$EU = [(\text{compensation})^{0.5}] \text{ if she does not work hard}$$

$$EU = [(\text{compensation})^{0.5} - 1] \text{ if she works hard}$$

What portion x must be paid to the manager to ensure that she chooses to work hard? This new compensation package must be competitive with the 100 flat salary.

4. A firm used to have productive assets that generated an income stream with a present value (PV) of 3,000. However, a fire occurred, and most of those assets were destroyed. The remaining undamaged assets produce an income stream that has a present value of only 1,000. Therefore the fire has reduced the value of the firm from 3,000 to 1,000. The firm could reconstruct the damaged assets for a capital cost of 1,500, which would restore the income

stream to its pre-loss level ($PV = 3,000$). The firm has existing debt of 2,000, which is a senior claim. Would the shareholders choose to reinvest by issuing new equity to pay for the loss, or are they better off walking away from the firm? Would the decision made by the shareholders be in the best interests of the bondholders? In answering this question, remember that the shareholders have limited liability, and therefore the share value cannot be negative.

5. CareLess Industries has two divisions. Division 1 makes cleaning products, and the net worth of this division (PV of cash flows) is 500. Division 2 makes a chemical product. The net worth of division 2 is 300, absent any potential liability. However, there is a chance that division 2 could have a 700 liability for pollution damage. The potential victims have no contractual relationship with the firm. The probability of such a loss is $0.2/(1 + s)$, where s is the amount the firm spends on safety. The firm must choose the level of s . If you could sell off division 2, would you do so? What is the gain from splitting the firm in this way? Assume a separated division 2 (as a stand-alone firm) is protected by limited liability. Note also that the derivative of $a/(1 + s)$ with respect to s is $-a/(1 + s)^2$; that is, $\Delta[a/(1 + s)]/\Delta s = -a/(1 + s)^2$.
6. SubAquatics (SA) sells scuba diving equipment. Its clients typically read specialist journals and are well informed about the price, reliability, and safety of SA and competitors' products. SA has estimated that, of 100,000 units sold each year at a price of \$100 each, there are $4/(1 - s)$ fatal accidents due to defective equipment. The value s is the amount spent by SA on safety in millions of dollars.
 - a. Assuming that SA is fully liable for such accidents and that the average settlement of each fatal accident is \$1 million, how much should SA spend on safety?
 Now assume that SA can escape this liability by selling its products at a lower price under a contract that allocates all responsibility for accidents to the purchaser (assume that courts enforce such contracts). If SA spends s (expressed in millions of dollars) on safety, the expected cost of accidents to any consumer is $[4/(1 - s)](\$1\text{m}/100,000) = \$40/(1 + s)$. Note that consumers are willing to pay \$100 when all liability is assumed by SA (assuming consumers are risk-neutral).
 - b. How much would consumers be willing to pay when they bear the cost of accidents?
 - c. How much would SA spend on safety?
 - d. Assuming that customers cannot observe the level of safety and there is no liability law, how much would SA spend on safety and how much would customers pay for the product?
7. A firm has existing operations that generate an earnings stream with a present value, PV , of 300 or 600, each with 0.5 probability. The firm has 250 in existing

debt. The firm wishes to undertake one of the following mutually exclusive new investments:

	Capital Cost	PV of Earnings	NPV
Project A	400	420	+20
Project B	400	0; probability 0.5 or 700; probability 0.5	−50

The capital cost of each project (400) is financed with new junior debt (face value 400). Is there an asset substitution problem? (Will shareholders try to choose the lower NPV project?) Show whether any asset substitution problem would disappear if the new project were financed with an equity issue of 400 instead of new debt.

EXCEL EXERCISE: MORAL HAZARD

Suppose that you own a building worth \$8,000,000 and have \$1,000,000 in cash. You own the building outright. Thus, your initial wealth (W) is \$9,000,000. Your utility (U) function is $U = W^{0.5}$.

There is a danger that your building will catch fire and burn down. We'll assume that fires are all or nothing events—that is, either they don't occur or, if they do occur, they result in a total loss of the building. Since you own the building outright, you are not required to buy fire insurance (as a mortgage holder would require of you). If you don't buy fire insurance, it is said that you self-insure.

You can also practice fire prevention (at a cost to you of \$150,000) or not practice fire prevention (at the cost of \$0 to you). If you practice fire prevention, the probability of a fire is 0.01. If you do not practice fire prevention, the probability of a fire is 0.05.

In a market with no insurance companies, you will have to self-insure. The question for you is whether you will practice fire prevention or not. Let's see.

Enter $=0.95*9000000^{0.5}+0.05*(9000000-8000000)^{0.5}$ in cell A1. This is your expected utility if you self-insure and don't practice fire prevention.

Enter $=0.99*(9000000-150000)^{0.5}+0.01*(9000000-150000-8000000)^{0.5}$ in cell A2. This is your expected utility if you self-insure and practice fire prevention. You should note that cell A2 exceeds cell A1, so we conclude that you will practice fire prevention in a world without insurance companies.

Suppose that an insurance company enters the scene and offers you a full coverage policy for price A5. This policy will cover all your damages (\$8,000,000) if a fire occurs.

Enter $=0.95*(9000000-A5)^{0.5}+0.05*(9000000-A5-8000000+8000000)^{0.5}$ in cell A3. This is your expected utility of buying a full coverage policy and not practicing fire prevention.

Enter $=0.99*(9000000-150000-A5)^{0.5}+0.01*(9000000-150000-A5-8000000+8000000)^{0.5}$ in cell A4. This is your expected utility of buying a full coverage policy and practicing fire protection.

Then play around with various values in cell A5. You will see that for any value of A5, cell A3 exceeds cell A4; in other words, for any full coverage insurance premium, you will not practice fire prevention (because your expected utility by not practicing fire prevention exceeds your expected utility from practicing fire prevention). You perhaps noticed that the formula you entered in cell A3 simplifies to $(9000000-A5)^{0.5}$, and the formula you entered in cell A4 simplifies to $(8850000-A5)^{0.5}$. Now it's easy to see that A3 is always greater than A4.

This is the moral hazard problem in insurance. Once the item is fully insured, the damage is no longer yours to bear, as it was when you self-insured; now the insurance company must bear it. The result is a tendency to not practice fire prevention (because it's expensive to do so). In addition, it is difficult and expensive for the insurance company to monitor the company's behavior (that is, to ascertain whether or not they practice fire prevention).

How much would you be willing to pay for full-coverage insurance? You'd have to get the same utility as you would get by self-insuring and practicing fire prevention, because that's the highest utility you can get if you self-insure, and you are under no obligation to insure. So if the insurance company wants you as a customer, they've got to beat or equal the best deal you currently have. We already know from above that if a full-coverage policy is offered, you will not practice fire prevention.

So enter $=0.95*(9000000-A7)^{0.5}+0.05*(9000000-A7-8000000+8000000)^{0.5}$ in cell A6. Note that this simplifies to $(9000000-A7)^{0.5}$. To find the maximum A7 you would pay, you would set $(9000000-A7)^{0.5}$ equal to A2. If you square both sides, you'll have $(9000000-A7)=A2^2$.

Enter $=9000000-A2^2$ in cell A7. This will give you the maximum you will pay for full-coverage insurance.

We will assume that our insurance company is risk-neutral. It can be risk-neutral because of the law of large numbers, which says that if it has a large number of customers like you, and if the fires are independent events, that is, my fire doesn't depend on your fire, and if the chance of a fire is 5%, then for every 100 customers, the insurance company can expect about 5 fires per time period. Although the insurance company doesn't know who will have a fire, it can count on about 5 fires happening. Because of this predictability, they can be risk-neutral.

Enter $=0.95*0+0.05*8000000$ in cell A8. This is the expected claim payout of the insurance company to you if you do not practice fire prevention.

Enter $=0.99*0+0.01*8000000$ in cell A9. This is the expected claim payout of the insurance company to you if you do practice fire prevention.

Enter $=A7-A9$ in cell A10. This is the insurance company's expected profit if it sells you a full-coverage policy and you practice fire prevention.

Enter $=A7-A8$ in cell A11. This is the insurance company's expected profit if it sells you a full-coverage policy and you do not practice fire prevention. This shows a loss, and since the insurance company knows that you will not practice fire prevention if you are fully insured, the company will not issue full coverage policies. If only the insured practiced fire prevention, the insurance company could offer a policy with an insurance premium of A7 (minus ϵ), and you would buy it (because you'd have a greater expected utility than if you practiced fire prevention and self-insured), and the insurance company would make a handsome profit (cell A10). But alas, because of the moral hazard problem, a profitable insurance market disappears.

However, there is a solution to the moral hazard problem. Because you (the insured) had no "skin in the game" after purchasing full coverage insurance—that is, none of the loss is on you—the market failed. Let's devise a policy that involves you having some skin in the game. Suppose the insurance company offers a policy for \$75,000 with a deductible of \$4,000,000 if the fire occurs; in other words, it will pay you \$4,000,000 of the \$8,000,000 damage you incur if there's a fire, but you incur the other \$4,000,000 in loss for a premium of \$75,000. Would you buy the policy?

Enter $=0.95*(9000000-75000)^{0.5}+0.05*(9000000-75000-8000000+4000000)^{0.5}$ in cell A12. This is your expected utility if you buy the deductible policy and don't practice fire protection.

Enter $=0.99*(9000000-75000-150000)^{0.5}+0.01*(9000000-75000-150000-8000000+4000000)^{0.5}$ in cell A13. This is your expected utility if you buy the deductible policy and do practice fire protection. You will notice that cell A13 exceeds cell A12 and so it appears that the deductible policy works; that is, you will buy the policy, and you will practice fire prevention.

But two things remain. You still have the option of not buying insurance. If you self-insure, we showed above that you would practice fire prevention. If we compare A13 with A2, we see that A13 exceeds A2, and so you are better off buying the deductible policy rather than self-insuring.

But will the insurance company want to sell it to you? Recall above that it didn't want to sell the full-coverage policy to you even though you want to buy it.

Enter $0.99*0+0.01*4000000$ in cell A14. This is the insurance company's expected payout to you with the deductible policy when you practice fire prevention (which is the utility maximizing thing for you to do).

Enter $=75000-A14$ in cell A15. This shows the insurance company's expected profit from selling the deductible policy to you when you practice fire prevention. This profit is \$35,000 if you've entered everything correctly.

We now have an "incentive-compatible" solution. The deductible policy allows the insured to increase their utility by buying the deductible policy, rather than self-insuring, and allows the insurance company to make a profit selling insurance. It's a win-win situation.

ADVERSE SELECTION

One factor that often affects managerial performance is how quickly managers learn about relevant information. Consider the Internet. Managers can achieve almost immediate access to information about rivals' products, and sellers can reach millions of potential customers. If you wish to buy an antique clock, of which only a handful may be for sale across the globe, you have a good chance of finding one of those rare sellers in the vast electronic world of eBay. If you wish to sell a used car, you want to expand the number of potential buyers to find one with a particularly high reservation price; the Internet gives you such wide access. This unprecedented flow of information profoundly affects markets. Monopolies tend to break down as consumers can search and compare across many sellers. The geographical boundaries of markets are expanding and some markets are truly global. Information fosters competition.

But information has more subtle effects on managerial behavior. We consider some of these in this chapter. Consider the differences in information between buyers and sellers. In many transactions, such as selling a car, securing a mortgage, buying health insurance, or investing in a company's stock, the buyer and seller have different information. For example, a seller of a used car usually knows more about the quality of the car than a buyer does. A borrower often knows more about his or her credit risk than does a lender. A policyholder knows more about his or her state of health than an insurer does. And "insiders" in a firm issuing shares of stock know more about the firm's prospects than the investors who

LEARNING OBJECTIVES

The Market for "Lemons"

Adverse Selection in Automobile Insurance

The Market for Annuities

Resolving Adverse Selection through Self-Selection

Using Education as a Signal: Adverse Selection in the Job Market

Using Warranties as Signals: Adverse Selection in the Product Market

Summary

Problems

Excel Exercise: Adverse Selection

may buy the stock. In these transactions, some individuals have more information than others. This chapter shows how managers can use their informational advantage to increase performance. And when managers are at an informational disadvantage, we illustrate how they can create creative defenses to mitigate their disadvantage.

THE MARKET FOR “LEMONS”

Some years ago the Nobel Prize-winning economist George Akerlof wrote a famous paper about the market for “lemons.” A “lemon” is a used car that turns out to have many faults not apparent at the time of sale—hence the sour taste. Let us call used cars that are virtually free of hidden defects “gems.” The fact that some cars are lemons and some are gems may simply be random. But what Akerlof had in mind was something a little more disturbing. A systematic process may ensure that a disproportionate number of lemons turn up in the used car market. This process arises from information differences between buyers and sellers. The basic idea is simple. Sellers know more about the hidden qualities of the cars they are selling than do buyers. If I have been driving a car, I know its defects; I know its mechanical record and whether it has been involved in accidents. Therefore, I know whether I am selling a lemon or a gem. The buyer can invest some time inspecting the car but is never going to be as well informed as the seller. The hidden qualities, good or bad, remain hidden.

Let us try to get into the mind of the buyer. How much is she willing to pay for a used car? She knows some cars are worse than average and some are better than average, but she does not know where the particular car she considers buying lies on this spectrum. So it seems safe to assume that she is willing to pay, at most, the value of a car of average quality. If the seller has a car that is a gem, he will be unwilling to sell at this average price because, known only to him, the car is really worth more than average. On the contrary, the seller of a lemon is delighted. He has a worse-than-average car and a buyer who is willing to pay a price based on average quality. If he does not blow the sale by appearing too eager, he can get a great deal. Therefore, at this average price, people tend to hold on to their gems, and the cars offered for sale are predominantly worse than average; they are lemons. And it is all because buyers are uninformed.

The picture gets worse. Buyers may not know whether a particular car is good or bad, but they can figure out what is happening. They can reason that only owners of lemons would offer their cars for sale at a price reflecting the average quality. Therefore, the selection of cars coming onto the secondhand market is not a true reflection of the overall population of cars but is mostly lemons. So buyers are not willing to offer a price that reflects the average quality. Indeed, because they can anticipate that only lemons will be sold, they are willing to buy only at a price appropriate to a lemon. Consequently no high-quality cars are offered for sale,

PROBLEM SOLVED: Equilibrium in the Used Car Market

Of course there are many types of used cars, so let us narrow things down to the market for 2006 Toyota Camrys. Some are better than average; and *if consumers knew they were buying one of these gems*, they would be willing to pay \$10,000 for it. Others are lemons; *if consumers knew they were buying a lemon*, they would be willing to offer only \$5,000. The problem is that consumers do not know which car is a lemon and which a gem. So they are willing to pay a price of \$7,500 reflecting the average quality.^a Sellers, of course, know the quality of what they are selling.

At a price of \$7,500

1. Owners of gems that really are worth \$10,000 would not sell for \$7,500.
2. Owners of lemons that are worth \$5,000 would be happy to sell for \$7,500.

Therefore, only the lemons will be sold. Now consumers, even though they do not know whether any particular car is a lemon or gem, anticipate that only the owners of lemons will choose to sell. Thus buyers assume that all cars being offered are lemons and are willing to pay only \$5,000. The lemon owners should still be willing to sell at \$5,000 because they know this is what their cars are worth.

Is the market described here really that simple? Of course we oversimplify things somewhat. Buyers are not totally ignorant, and sellers do not know everything. For example, some sellers really do not know the quality of their cars, and it is possible that some

owners of high-quality vehicles will sell at the average price. Other sellers may believe their cars to be high-quality but are willing to sell because they have an urgent need for money. Some buyers may believe they have better information or a nose for a good deal. The market is not as simple as we portrayed here, and some high-quality cars are sold in the second-hand market. But the basic ideas of adverse selection still hold in this more realistic world. If it is generally true that sellers on average are more informed than buyers, there may be some dispersion in secondhand prices, but these prices tend to converge toward the average price. Consequently, while a few people may sell gems at this price, the cars offered for sale are mostly lemons. So we still get a used car market stocked predominantly with lemons and relatively few high-quality cars.

We can now see why the term adverse selection is used. Where buyers are unable to distinguish quality, the price averages across quality groups offered for sale. However, at this common price, the selection of cars being offered for sale is not representative; rather it is weighted toward the low quality—it is adverse.

^aWe assume there are equal numbers of lemons and gems. For example, with 100 lemons and 100 gems, the average price is $[100(\$10,000) + 100(\$5,000)]/200 = \$7,500$. If there were different numbers, then the average price would differ. For example, with 150 gems and 50 lemons, $[150(\$10,000) + 50(\$5,000)]/200 = \$8,750$.

only lemons; and the price reflects this low quality. The market for high-quality used cars has disappeared.

You might object to this analysis. Surely the seller of a high-quality car can tell the buyer, “My car is better than average, and you should pay a high price.” The problem is that this statement is not credible. The buyer cannot verify this statement, so the owners of lemons have every reason to declare that their cars also are wonderful. Cheap talk is simply not convincing because it can be mimicked by the

owners of lemons. Later we show that there are mechanisms to separate the gems from the lemons, but these involve more sophisticated signaling.

Economists often describe a lemons market as a market failure. It is certainly desirable from everybody's viewpoint to have a vigorous market for high-quality used cars. But due to asymmetric information, this market is stunted or killed completely. As we go through this chapter, we look at how uninformed managers can compensate for their ignorance and redevelop many of these damaged markets.

ADVERSE SELECTION IN AUTOMOBILE INSURANCE

Although the term lemons was first used for cars, the term adverse selection was first used in insurance. We look now at how adverse selection arises in various insurance markets. We first look at automobile insurance; after that we look at annuities and life insurance.

If managers at the insurance firm (the insurer) can distinguish drivers according to their respective loss characteristics, each policyholder can be charged a premium that precisely matches his or her expected loss. Good drivers pay low premiums, and bad drivers pay higher premiums. Insurers take some trouble to try to tailor premiums in this way. They ask questions about observable characteristics, such as automobile type, location, or age and gender of the policyholder. And by careful statistical analysis of their databases, they determine how each of these characteristics predicts accident rates.¹ This information is then used to set premiums. But even after classifying in this way, differences in risk may remain between policyholders. For example, the insurer may set a premium based on age, gender, vehicle type, and location. But not all 22-year-old men driving sedans in Philadelphia have the same loss potential. Their skill levels and behavioral characteristics can vary substantially. So there is an effective subsidy from low-risk drivers to high-risk drivers within each class. This subsidy can have an unsettling effect.

In Figure 16.1 we consider a category of drivers in the insurance pool—say the 22-year-old men driving sedans in Philadelphia. Some are worse drivers than others. We call these respectively *high-* and *low-risk drivers*. Each policyholder has a wealth level of 125, but a loss can reduce the wealth to 25; that is, drivers can lose 100 of their wealth should the loss occur. For the high-risk group, the probability of loss is 0.75, resulting in an expected loss of $0.75(100) = 75$. For the low-risk group, the probability of loss is 0.25, resulting in an expected loss of $0.25(100) = 25$.

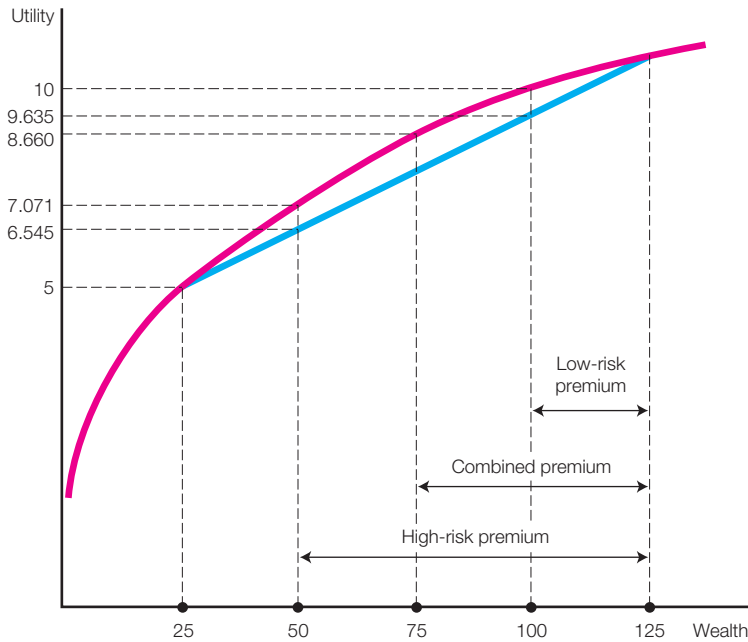
Perfect Information

First we show that, if the insurer could distinguish between the two groups, managers could charge competitive premiums, and each group would buy insurance. The competitive premiums for each group are their respective expected losses of

1. Note that we are not saying that features such as age and gender differences *cause* differences in accident rates—only that there may be a statistical association.

FIGURE 16.1

Adverse Selection in Automobile Insurance



75 and 25 (in practice, the insurer would add an allowance for transaction costs and profit). Using expected utility, we can now show that each person will buy insurance. We need a utility function that reflects risk aversion. (If people were not risk-averse, why would they buy insurance?)

$$\text{Utility} = (\text{Wealth})^{0.5}$$

For the low-risk group, the utility of insuring and having wealth of 100 with certainty (derived by subtracting the premium of 25 from the initial wealth of 125)—that is, $U(100)$ —is higher than the expected utility of not insuring (where wealth stays at the initial 125 without loss but falls by 100 to 25 if a loss occurs)

$$\text{Utility with insurance} = (125 - 25)^{0.5} = 10$$

$$\text{Utility with no insurance} = (0.75)(125)^{0.5} + (0.25)(25)^{0.5} = 9.635$$

Similarly, for the high-risk group,

$$\text{Utility with insurance} = (125 - 75)^{0.5} = 7.071$$

$$\text{Utility with no insurance} = (0.25)(125)^{0.5} + (0.75)(25)^{0.5} = 6.545$$

The respective positions are shown on the vertical axis of Figure 16.1.

STRATEGY SESSION: What Does Big Brother Know and When Does He Know It?

There's a new smartphone app that tracks and offers promotions to shoppers as they are outside the store (they could go inside to buy), as they move inside (they could buy and are more likely to than when outside), as they go into dressing rooms (people who use dressing rooms are more likely to buy than people just in the store), as they are at the cash register (they have made a purchase and may be enticed to buy more on this or a future visit). The app is called Shopkick and is compatible with iPhones and Androids. It is being used by Macy's, Best Buy, Sports Authority, American Eagle Outfitters, and mall operator Simon Property Group. Those are some pretty prominent retail names.

How does it work? You can get points by checking into a participating store. You don't have to go inside, just near enough to the store so that it can detect you're there. You get a higher number of points if you go in the store, a still higher number if you handle the merchandise, and an even higher number if you use a dressing room. Points are then redeemable for gift cards and other gifts. It takes a lot of points just to get a \$5 gift card, and the stores limit the number of times per day you'll get credit for a store visit, so don't think that you'll get rich by spending a day going in and out of stores.

If people use Shopkick, Shopkick will learn about your shopping locations, when you buy things, what you buy, what you don't buy, how much you spend, what you do to excess, and so on. Of course, they are doing this with your permission; if you have read

their agreement beyond the point where it says you could get a gift card. Stores could adjust points to highlight products as they wish. Knowing you are in the store, the staff can approach you or offer a special promotion or a discount at checkout. Everything can be tailored to the individual customer. The app can reward consumers instantaneously for desired behavior (desired from the merchant's point of view that is). As Mike Dupuis, vice president of marketing and operations at American Eagle Outfitters puts it, between the app and the company's reward card number, "We have the ability to target down to even an individual level."

Best Buy is doing the same. Rick Rommel, a senior vice president, states that discounts can be triggered immediately when a customer enters the store and tailored to that customer's specific interest. If the customer had browsed Best Buy for computers previously, he or she could be offered a discount on a computer. "This really moves toward one-to-one marketing," says Rommel. Other retailers want the ability to give points or base promotions on characteristics such as sex, age, residence, shopping frequency, buying history, and so on.

Source: Stephanie Clifford, "Aisle by Aisle, an App That Pushes Bargains," *New York Times*, August 17, 2010, at www.nytimes.com/2010/08/17/technology/17app.html; and Steve Lohr, "Innovate, Yes, but Make It Practical," *New York Times*, August 15, 2010, www.nytimes.com/2010/08/15/business/15unboxed.html.

Asymmetric Information

Now suppose managers at the insurer are unable to distinguish between high- and low-risk drivers. However, we assume that each driver knows whether he or she is a high or low risk; in effect we each know our own driving capabilities. If there are equal numbers of low- and high-risk drivers, the insurer can break even by charging the average premium of $0.5(25 + 75) = 50$. But will each group now

continue to buy insurance? It is clear that the high-risk drivers will buy insurance. If they would have bought insurance at a premium of 75, then surely they still will buy if the premium is reduced to 50

$$\text{Utility with insurance} = (125 - 50)^{0.5} = 8.660$$

$$\text{Utility with no insurance} = (0.25)(125)^{0.5} + (0.75)(25)^{0.5} = 6.545$$

But for the low-risk group, the insurance premium has increased from 25 to 50. We can compare the expected utility with and without insurance for this group

$$\text{Utility with insurance} = (125 - 50)^{0.5} = 8.660$$

$$\text{Utility with no insurance} = (0.75)(125)^{0.5} + (25)^{0.5} = 9.635$$

So the low-risk group will not buy the insurance. The only people who will buy insurance at the common price of 50 are the high-risk drivers. If you think a step ahead, you can see that this pricing structure is not tenable for the insurer. If only high-risk drivers buy the insurance, then each policyholder costs 75 (on average) in claims but pays a premium of 50. The insurer loses 25 on average for each policy and cannot stay in business unless premiums are increased to 75. The insurer offers full insurance only at a premium of 75. Who will buy this policy? We saw that the high-risk drivers get utility of 7.071 from buying this policy but have expected utility of 6.545 if they buy no insurance. So high-risk drivers buy the insurance. Good drivers do not buy the policy as now shown

$$\text{Utility with insurance} = (125 - 75)^{0.5} = 7.071$$

$$\text{Utility with no insurance} = (0.75)(125)^{0.5} + (0.25)(25)^{0.5} = 9.365$$

This illustration is a perfect lemons market. Because managers of the insurer cannot perfectly discriminate between high- and low-risk drivers, they offer only one type of policy, which appeals to only the high-risk drivers. An adverse selection of drivers choose to purchase insurance. The people who suffer from the information problem are the low-risk drivers, who are priced out of the market.

Managers can help restore the market in two ways so that the low-risk drivers have an appropriate choice of policies. Competition between insurers may help reduce the problem. Information about loss expectancies of individual drivers is of economic value to an insurer, so there is an incentive for managers at each insurer to invest in better information about policyholders. If an insurer acquires better information about policyholders than a rival does, that insurer can selectively attract low-risk drivers from the rival who is unable to discriminate simply by offering a lower price and admitting only low-risk drivers. Thus competition induces insurers to seek and compile information that lets them use premium structures to discriminate among risk groups. Of course insurers will never find out everything they need to know about their policyholders, and adverse selection never disappears. But in an actively competitive market, adverse selection is reduced to a level that reflects the cost of information.

STRATEGY SESSION: Adverse Selection and the National Health Care Debate: The Federal Mandate on Compulsory Insurance

Federal legislation signed in 2010 requires Americans (with some hardship exceptions) to carry health insurance by 2014. The issue is highly controversial and its constitutionality is being challenged in the courts. The case for mandating health insurance appeals to adverse selection

If insurance is not mandated, then some will buy it and some will not. Those in good health are less likely to buy, and those in poor health are more likely to buy. This adverse selection will be intensified insofar as insurers are forbidden to charge premiums that reflect individual risk. The lack of risk-related premiums implies that those in good health will subsidize those in poor health. But if those in good health do not buy insurance, the subsidy is withdrawn and premiums will have to rise. The Congressional Budget Office (CBO) has estimated that, even with the mandate, there would be 20 million uninsured. But, further estimates by the Urban Institute suggest that the uninsured would swell by a further 17.8 million people if the mandate were not adopted. Taking out these 17.8 million would then cause premiums to rise by 15–29% according to the CBO.

It is interesting to see moral hazard figures in this issue. We saw in Chapter 15, that when people

are insured, they tend to pay less attention to costs. The experience in Massachusetts makes critics of mandates nervous. A 2005 law in that state offered subsidized insurance to those with low incomes. Many uninsured poor were making visits to the emergency room (often for routine ailments) and the intention was to replace the 28% unpaid hospital bills with a 28% premium subsidy. But emergency visits did not fall; in fact, they rose as did health care costs. Thus, Massachusetts, already the second highest state in health care costs, found it costs rising even further.

Thus, we see an interesting tension here between adverse selection and moral hazard.

Sources: Rahul Rajkumar and Harold Pollack. (January 7, 2011) "A Mandate That Matters; Requiring Health Insurance Supports the Private Market and Lessens Government's Role," *Los Angeles Times*; Liz Kowalczyk (April 24, 2009). "ER Visits, Costs in Mass. Climb," *The Boston Globe*. http://www.boston.com/news/local/massachusetts/articles/2009/04/24/er_visits_costs_in_mass_climb/; Kay Lazar, (August 22, 2009). "Bay State Health Insurance Premiums Highest in Country," *The Boston Globe*. http://www.boston.com/news/health/articles/2009/08/22/bay_state_health_insurance_premiums_highest_in_country/.

The second way in which managers can help restore markets is through strategic design. We discuss this later, but we offer a teaser here. Is there any way to induce policyholders to reveal information about themselves in a credible way? Suppose managers offer a choice of A or B to each driver. Can they design these alternatives so that only a low-risk driver will choose A and only a high-risk driver will choose B? If managers can develop such a menu, they create a separating equilibrium, which we will discuss later in this chapter. Only low-risk drivers choose A, and only high-risk drivers choose B. Managers can infer what risk type each person is by the menu item he or she selects. This idea of self-selecting menus is discussed shortly.

THE MARKET FOR ANNUITIES

Another interesting adverse selection illustration is the market for annuities. Many managers have defined contribution retirement plans. Under these plans, the employer or the employee (or both) make explicit contributions to a pension plan. The money is invested in some investment vehicle, and the value of the invested assets at retirement is available to the retiree. There is no guarantee of an income in retirement (unlike defined benefit plans). Suppose employees want to take this sum of money and convert it into an income stream that will last until they die. For example, if I knew I would live for 10 years, I could take my cash, divide it by 10 (with a little adjustment for investment income), and have a constant income stream for the rest of my life. The problem is that I cannot predict when I will die. In this savings scheme, I run out of money if I live for more than 10 years.

An annuity helps solve my problem. Even though I do not know how long I am going to live, I can buy an annuity that converts my principal into a constant income stream for the remainder of my life. Essentially, I give my cash to an annuity firm (usually a life insurance company), which in return promises to pay me a constant annual sum for as long as I live. Thus I insure against living too long and running out of money.

What are the incentives of the managers at the annuity firm? They sell many annuities. But consider 1,000 65-year-old women, all of whom buy annuities. We will work through two simple examples of how annuity markets work. In the first case the health status of each individual is known both to that person and to managers at the annuity firm—perfect information. In the second case there is asymmetric information—the person knows her health status but the managers do not.

The Cost of Information Asymmetries in Annuity Markets

Annuity Markets with Full Information

Let us consider an annuity market in which managers have perfect information. Our 65-year-old women differ. All are not in the same state of health, and consequently they have different life expectancies:

- A quarter of our population is in poor health and is expected to live for only 5 years.
- Half are in average health and have a life expectancy of 15 years.
- A further quarter is in excellent health, with a life expectancy of 25 years.

Each person has \$300,000 in capital and wishes to buy an annuity. Because the annuity firm knows each person's life expectancy, it can make the following deals

STRATEGY SESSION: Information Asymmetries and Brownfields

Have you ever wondered (and worried) about those abandoned and contaminated industrial sites or gas stations that linger for years, often without being cleaned up or sold and put to use? Part of the problem turns out to be adverse selection.

The background to this problem is found in legislation enacted in 1980. The Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA), commonly known as Superfund, was passed to make polluters responsible for contaminated land. Being made liable, the polluters would be responsible for the cost of cleaning up, which in many cases could be enormous. The parties who had this burden thrust upon them included not only those who put toxic materials on the land but also the owners of the land. Moreover, the liability for cleanup could pass to a new owner when the property was sold, even if the new owner did not contribute to the contamination. Does this impede the sale of such land? After all, the price that the land sells for should be discounted by the cost of cleanup, thus making it attractive for buyers to acquire and redevelop the land. However, two types of asymmetries can arise to stymie such sales.

The first of these is the so-called judgment-proof problem. If you are sued for cleanup, the amount you must pay is limited by your net worth (or, if you are a firm, by the equity value of the firm). Suppose a sale were to take place in which the net worth of the seller was considerably less than the net worth of the buyer (which might be likely because the seller's business has declined to the point that the land has been abandoned). Consider that the cost of cleanup could be as much as \$10 million, and there is a 50% chance the owner will be asked to clean up. The seller's net worth is \$5 million, and the buyer is worth \$20 million.

Because the seller could never have paid more than her net worth, this means the seller will have a 50% chance of escaping a \$5 million liability and will therefore be willing to reduce the price by only 0.5 times \$5 million = \$2.5 million. However, the buyer will assume a 50% chance of paying the full \$10 million and will want the price reduced by \$5 million. You can easily see why the buyer may not agree on a price: The seller is holding out for a price the buyer would be unwilling to pay, and therefore the sale may not take place, and the land will continue to lie idle.

But there is another problem. The seller knows more about the property and therefore the potential liability than the buyer. Uninformed buyers therefore are likely to be pessimistic about the likelihood and cost of cleanup and will therefore want a large price discount. At this discount, only sellers whose land is very badly contaminated are likely to sell. This creates the familiar adverse selection problem similar to the others addressed in this chapter: Only the most contaminated land will sell. In fact, even the fear that land might be contaminated may lead buyers to assume the worst, driving high-quality land from the market. But buyers can undertake an environmental audit to establish how polluted land really is. Although this may redress the information balance, the costs of such audits are high. Thus even the idea of paying for an audit because of the fear of adverse selection has a big deterrent effect on land sales.

Source: Derek B. Singer, "Brownfield Remediation as a Policy Tool in Urban Redevelopment," Working Paper EC465, Department of Economics, Middlebury College, November 2005.

(to keep things simple, we ignore interest and inflation and allow the annuity firm to break even):

- Those with a life expectancy of five years are offered an annuity of \$60,000 per year. The firm receives \$300,000 up front and pays five annuity payments of $\$60,000 = \$300,000$.
- Those with a life expectancy of 15 years are offered an annuity of \$20,000 per year. The firm receives \$300,000 up front and pays 15 annuity payments of $\$20,000 = \$300,000$.
- Those with a life expectancy of 25 years are offered an annuity of \$12,000 per year. The firm receives \$300,000 up front and pays 25 annuity payments of $\$12,000 = \$300,000$.

In an economic sense, this is a nondiscrimination case. Each person receives an annuity based on her own health status, and no one subsidizes anyone else. Moreover, if annuity markets were perfectly competitive, this is how managers would behave.

Annuity Markets with Asymmetric Information: Adverse Selection

Now assume the annuity firm charges everyone a premium of \$300,000 and pays each person \$20,000 per year until her death. This same annuity can be purchased by anyone. To see what adverse selection does to behavior in this market, we ask who knows what about each person's health. Managers of annuity firms can often find out something about health; they can ask for a medical test.² But what are the managers worried about when they ask for a test? They are certainly not worried about people being ill and having a five-year life expectancy. For each of these people, the firm receives \$300,000 and pays five installments of \$20,000, thus clearing a profit of \$200,000. On the contrary, they are worried about people being "too healthy" and living too long. The annuity firm loses \$200,000 on anyone who lives 25 years (it receives \$300,000 and pays 25 installments of \$20,000). In testing people, the managers try to spot those who are healthy and will live a long time. This is not easy. Verifying that people are ill is one thing. Identifying people who may wish to pretend they are ill but are indeed healthy is quite difficult. Finding health problems is easier than finding the absence of them. So the managers often lack reliable information about the health status of customers. On the other hand, each person knows a lot about his or her own health status: medical history, dietary and exercise habits, and so forth. So we have the classic information asymmetry that can lead to adverse selection.

Because managers do not know the health status of any particular applicant, they estimate the firm can break even. To see this, the total income for the firm is

$$1,000 \times \$300,000 = \$300,000,000$$

2. In practice, the annuity sum reflects some markup for the firm's profit and also accounts for investment income.

Here are the total annuity payments

250 people ($\frac{1}{4}$ of 1,000) die after 5 years

$$= 250 \text{ times } \$20,000 \times 5 \text{ years} = \$25,000,000$$

500 people ($\frac{1}{2}$ of 1,000) die after 15 years

$$= 500 \text{ times } \$20,000 \times 15 \text{ years} = \$150,000,000$$

250 people ($\frac{1}{4}$ of 1,000) die after 25 years

$$= 250 \text{ times } \$20,000 \times 25 \text{ years} = \$125,000,000$$

$$\text{Total payments} = \$300,000,000$$

Another way to see this example is to look at the after-the-fact income statement for each group.

- Those dying after five years pay \$300,000 for the annuity and receive $(5)(\$20,000) = 100,000$. The annuity firm makes a profit of \$200,000 on each of these individuals.
- Those dying after 15 years pay \$300,000 for the annuity and receive $(15)(\$20,000) = \$300,000$. The annuity firm breaks even on each of these individuals.
- Those dying after 25 years pay \$300,000 for the annuity and receive $(25)(\$20,000) = \$500,000$. The annuity firm loses \$200,000 on each of these individuals.

STRATEGY SESSION: Adverse Selection in the Federal Government Prescription Drug Plan

Adverse selection can bedevil both private and public programs. Consider the prescription drug plan (PDP) created by the U.S. Medicare Modernization Act of 2003. Since 2006 the PDP, as the name suggests, has provided coverage for prescription drugs. The issue arises because the plan is optional. Recent estimates have attempted to see whether enrollment will be dominated by those most likely to use the plan. Research reported at a recent meeting of health economists suggests that this is indeed the case. Excluding those enrolled in Medicare and current employer-sponsored plans, "... approximately 18% of the remaining beneficiaries will enroll in a PDP without an additional supplement. Drug expenditures for this group will be about 11% higher than the popu-

lation average, indicating adverse selection. A similar number are predicted to enroll in a nondrug Medigap plan as well as a PDP (approximately 18%). These beneficiaries will have drug expenditures typically about 21% higher than the population average, and beneficiaries enrolling in FFS and nondrug Medigap (4%) will have average drug expenditures about 31% higher. By contrast, HMO enrollees (36%) and enrollees in FFS only (24%) will have drug expenditures about 12% lower than the population average."

Source: S. Pizer, A. Frakt, and R. Feldman, "Storm Clouds on the Horizon—Expected Adverse Selection in Medicare Prescription Drug Plans." Paper presented at the annual meeting of the Economics of Population Health: Inaugural Conference of the American Society of Health Economists, Madison, WI, June 4, 2006.

Who buys the annuity? If those in poor health believe they have a life expectancy of only five years, then five annual payments of \$20,000 at an up-front cost of \$300,000 are a bad deal. These people would probably be much better off simply drawing down their capital. For example, if they allow themselves \$30,000 a year, they will not run out of money until after 10 years (twice their life expectancy). So those in poor health are unlikely to buy the annuity paying \$20,000. Those in average health may be inclined to buy because they remove some uncertainty and get a reasonable financial deal (remember that we are ignoring interest). Those in excellent health find the annuity of \$20,000 per year to be a fabulous deal—they pay \$300,000 and get back \$500,000. So this deal appeals to only some of our 1,000 65-year-olds. Only those in average or better-than-average health will buy the annuity.

Of course managers can anticipate that only those in average or better-than-average health will buy the policy; so if it offers an annuity of \$20,000, the firm will lose money. To see this, note that managers will break even on each of the 500 people in average health and lose \$200,000 on each of the people in excellent health. Therefore, they have to reduce the annuity's value until they achieve break-even costs on those choosing to buy at this reduced value.

Evidence of Adverse Selection in the Annuity and Life Insurance Markets

Managers can use a simple test to see whether annuity markets are subject to adverse selection. Recall that in the preceding illustration, the population held 1,000 people and the average life expectancy was 15 years, calculated as follows

$$\frac{1}{4} (5 \text{ years}) + \frac{1}{2} (15 \text{ years}) + \frac{1}{4} (25 \text{ years}) = 15 \text{ years}$$

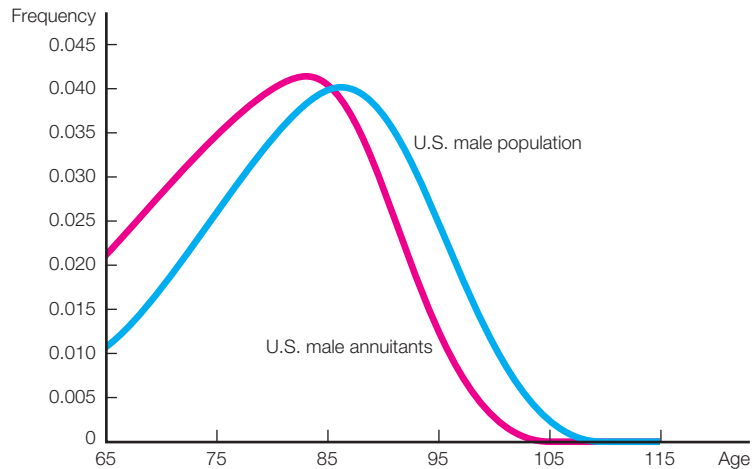
If the annuity was offered at \$20,000, then only the 750 people in average or better-than-average health would actually buy it. Therefore, the life expectancy of the annuitants would be

$$(500/750)(15 \text{ years}) + (250/750)(25 \text{ years}) = 18.333 \text{ years}$$

If managers could observe life expectancy, they could see whether it was greater for the annuitants than for the population as a whole. The problem is that they cannot observe life expectancy. However, they can observe the actual mortality rates of populations. If the average life expectancy of annuitants is indeed higher than that for the population as a whole, managers will find that those buying annuities live, on average, longer than the population as a whole. We draw on some evidence from two colleagues at Wharton comparing the life span of annuitants with the population in the United States.³ Figure 16.2 shows the distribution of age of death for those buying annuities and for the whole U.S. male population. The distribution for annuitants is clearly shifted to the right, indicating that they indeed live longer on average.

3. This is taken from David McCarthy and Olivia Mitchell, "International Adverse Selection in Life Insurance and Annuities," 2003.

FIGURE 16.2

Mortality Distributions for the U.S. Population and Annuitants

Source: O.S. Mitchell and D. McCarthy. "Annuities for an Aging World," in E. Fornero and E. Luciano, eds., *Developing an Annuities Market in Europe* (Northampton, MA: Edward Elgar, 2003).

The Absence of Adverse Selection in Life Insurance

Whereas annuities insure people against “living too long” and running out of money, life insurance protects the survivors of people who “die too soon.” If managers at life insurance firms have less information about the health status of their policyholders than the policyholders themselves, we would expect adverse selection here too. But do managers indeed have less information?

Recall that when we discussed annuities, we suggested that it might be challenging for managers to exclude those in good health; establishing the absence of poor health is difficult. The evidence on annuities showed that some degree of information asymmetry exists.

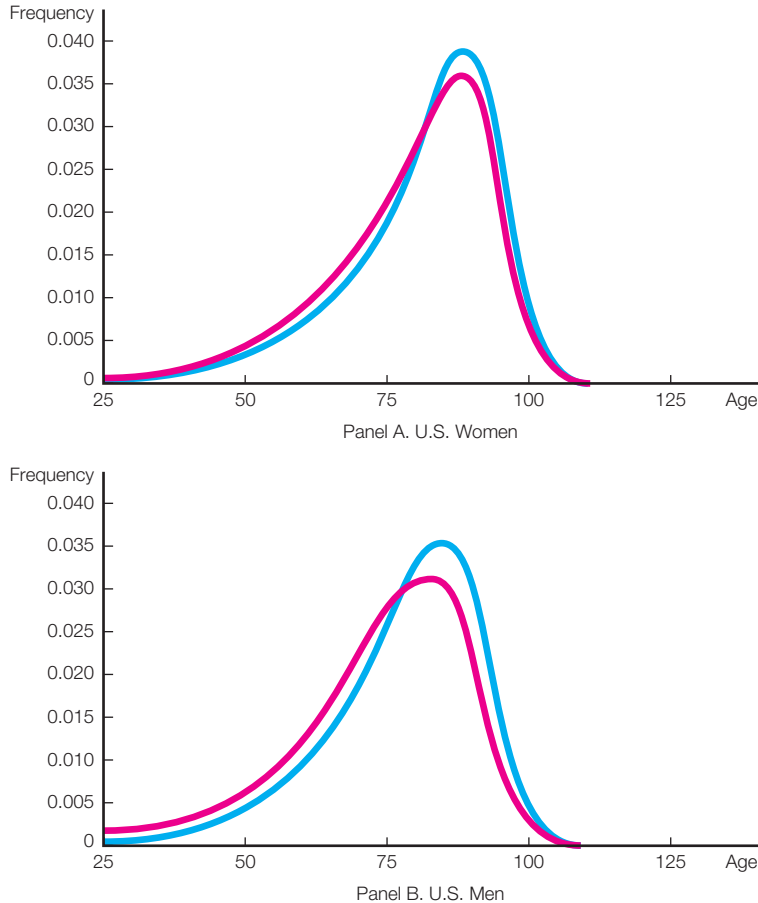
Managers at a life insurance firm routinely conduct medical examinations of those seeking life insurance and turn down (or charge higher premiums to) those who do not test well. Does this action by managers significantly reduce the information asymmetries and allow them to adequately predict future health? Let us see.

In Figure 16.3 we show some evidence drawn from the United States for both men and women. These figures show the histograms of mortality—that is, how many people die in different age groups. If insurance companies did not conduct effective medical exams and there was adverse selection, we would expect that healthier people would be less inclined to buy insurance and that the mortality rates for the insured population would be higher than for the population as a whole. But this is not what we see. The distribution for the insured population is

FIGURE 16.3

Distribution of Age at Death Conditional on Attaining Age 25

The red line shows the distribution of age of death for the population, and the blue line shows the distribution of age of death for those purchasing life insurance.



Source: David McCarthy & Olivia S. Mitchell, "International Adverse Selection in Life Insurance and Annuities." In *Riding the Age Waves: Responses to Aging in Advanced Industrial States*, edited by Shripad Tuljapurkar, Naohiro Ogawa, and Anne Gauthier. Elsevier. *Forthcoming*.

clearly shifted to the right, indicating that mortality rates are lower among those who have life insurance. This pattern can be seen for both men and women, but it is less dramatic for women. The same pattern has been observed in the United Kingdom. It appears that, far from having an information disadvantage, managers at life insurance firms have been very effective in establishing the health status of their policyholders and offering insurance predominantly to those in good health.⁴ There seems to be no traceable adverse selection here.

4. There is another possible explanation for these results. People who buy life insurance probably have above-average wealth. And wealth is also associated with health and longevity. Therefore, the reason that insured people live longer may have more to do with their wealth rather than insurance company screening.

RESOLVING ADVERSE SELECTION THROUGH SELF-SELECTION

While Akerlof described the adverse selection problem, Michael Rothschild and Joseph Stiglitz laid out an elegant solution to adverse selection. (Stiglitz shared the 2001 Nobel Prize with Akerlof and Michael Spence.) The idea behind their solution is simple. If the buyer of a secondhand car simply asks about quality, the answer is not credible. If an insurance firm simply asks how good a driver you are, you are certainly going to embellish your answer. Asking is not enough. How can an uninformed manager obtain credible information? Consider, for example, managers at the insurance company. They do not know whether any driver is a high or low risk. However, the insurer does know the following:

1. Some drivers are high risk and some are low risk.
2. The drivers themselves know whether they are good or bad drivers.

Managers should reason that, because individuals know their own risk type, they might sometimes use this private information to reach different decisions. So managers should design policies that separate behavior between good and bad drivers. Good drivers choose one policy and bad drivers choose an alternative policy. Managers induce drivers to reveal their risk types by the policies they choose.

Let us think of a choice problem that does the trick. Managers might offer every driver who seeks insurance a choice between two policies—full insurance or a high deductible:

Full insurance When every loss is paid in full.

Deductible When the insurer does not pay the full loss but pays the loss minus some fixed amount.

- Policy A has a high premium (designed to break even if bought only by high-risk drivers) and offers **full insurance**. Full insurance means that every loss is paid in full.
- Policy B has a much lower premium but a big **deductible**. A deductible means that the insurer does not pay the full loss but pays the loss minus some fixed amount. For example, suppose the policy has a \$2,000 deductible. If there is a \$20,000 loss, the insurer will pay $\$20,000 - \$2,000 = \$18,000$. With a \$5,000 loss the insurer pays $\$5,000 - \$2,000 = \$3,000$, and so forth. If the loss is less than \$2,000, the insurance firm pays nothing.

Drivers who know they are bad are likely to have one or more claims, so the deductible is a big deterrent. It is much better for bad drivers to pay the higher premium and avoid the deductible. But drivers who know they are good will reason the premium saving is more important because they are unlikely to have a claim and therefore are unlikely to face the deductible. So good drivers select the cheaper deductible policy and bad drivers select the more expensive, full coverage.

Managers can also use flat and experience-related premiums to get customers to reveal their true type. Consider the following:

- Policy C has a relatively high price. Furthermore, the driver can buy the policy next year, and the year after, and so on at the same price. The premium does not go up (except for inflation) even if the driver has claims under the policy.
- Policy D starts with a high premium; in the first year the premium for policy D is higher than for policy C. However, the premium for policy D changes in future years according to the number of accidents the driver has. If there are no claims, the premium for policy D falls to a level far below that for policy C. If there are claims, the premium stays at the high level—that is, a little higher than policy C.

Now consider the reasoning of those who drive badly. They can choose either policy. But they know they are bad drivers, and there is a high chance they will crash their cars and make an insurance claim. Therefore, policy D looks unattractive. The potential premium reduction for having no claims is not really relevant given their self-knowledge. They know if they buy policy D, they always (or at least often) pay a higher premium. But good drivers see the choice differently. For them, the chance of a big premium reduction is attractive because they believe they are unlikely to have a claim. Their likelihood of getting a premium reduction more than makes up for a slightly higher premium in the first year. Accordingly, good drivers will tend to choose policy D, and bad drivers will choose policy C.

Simple Adverse Selection

In our earlier example, each driver had an initial wealth of 125 and could lose 100 with a crash. The probabilities of a crash were

$$\begin{aligned}\text{Probability that bad drivers crash} &= 0.75 \\ \text{Probability that good drivers crash} &= 0.25\end{aligned}$$

We know from this example that the market for good drivers collapses altogether. Indeed, the only policy offered is a policy offering full insurance at a premium of $0.75(100) = 75$, designed to break even for high-risk drivers. Only bad drivers buy insurance at this price. No insurance policy is offered that would appeal to good drivers. Can managers do better?

Suppose the insurance company managers offer the following two policies, and each driver can choose which, if any, to buy:

- Policy 1 charges a premium of 75, but it fully pays for the loss of 100 if a crash occurs. Anyone buying this policy will have a wealth of 125 minus the premium of $75 = 50$. Note that because of the full insurance, the wealth of the individual is unaffected by whether the loss occurs.

- Policy 2 pays a fixed sum of 10 if a loss occurs, and the premium is 2.5. Anyone buying this policy will have the following wealth

With no loss: 125 minus the premium of 2.5 = 122.5

With a loss: 125 minus the premium of 2.5,
minus loss of 100, plus the payment of 10 = 32.5

Clearly, policy 2 is not a perfect insurance policy, but it offers some compensation for loss. Let us examine the choice facing each type of policyholder. First

STRATEGY SESSION: How Do Venture Capitalists Resolve Adverse Selection and Moral Hazard Problems? The German Experience

Adverse selection and moral hazard issues may explain what forms of financing are used by venture capital firms to fund portfolio firms.

A. The first issue is moral hazard. The entrepreneur has control over the investor's money. Will the entrepreneur act in the best interest of the investor? To resolve this issue generally calls for the entrepreneur to have some "skin in the game"; i.e., he or she should have some real equity stake in the game.

B. The second issue is adverse selection; how much does the investor know about the quality of the company in which it is investing. If there is poor information, the investor may also require that the entrepreneur share the risk through equity (or similar) financing. But with good information, it is less important that the entrepreneur signal his hidden information by sharing the risk—thus debt, and similar forms of financing, become acceptable.

Now, let's ask whether the particular form of financing is chosen to mitigate the moral hazard and adverse selection problems.

Broadly speaking, venture capitalists can finance with debt, equity, or some intermediate (mezzanine) financing that has some of the features of debt and some features of equity (examples are convertible debt,

preferred equity, and convertible preferred equity). The different forms of financing allocate risk in different ways. For example, straight debt removes risk (except default risk) from the investor; convertible debt¹ allocates upside risk (positive risk or gain) to the investor, but protects the investor from downside risk; and equity allocates unlimited upside and partial downside risk (investors can avail themselves of limited liability to default) to the investor.²

In Germany, a particular form of mezzanine financing is used—a silent partnership. This usually divides the payoff to the investor according to the return made by the firm.

1. If the return is low, the investor gets a fixed interest payment (thus offering the investor some downside protection and some decision rights). This payoff is junior to that of the conventional debt holders.
2. For intermediate levels of returns, the investor receives a claim on these cash flows—like an equity position but it is senior to that of the shareholders.
3. But there is a cap on the "equity-like" payoffs. Thus, for high returns, the investor gets a fixed payoff—like debt.

These silent partnerships are really part debt/part equity. By changing the parameters (the boundaries

look at the high-risk, bad drivers. Note that high-risk drivers (and low-risk drivers later) know their probability of loss, so they use this probability when calculating expected utility

$$\begin{aligned}\text{No insurance utility} &= (0.25)(125)^{0.5} + (0.75)(125 - 100)^{0.5} = 6.545 \\ \text{Policy 1 utility} &= (125 - 75)^{0.5} = 7.071 \\ \text{Policy 2 utility} &= (0.25)(125 - 2.5)^{0.5} - (0.75)(125 - 100 - 2.5 + 10)^{0.5} \\ &= 7.043\end{aligned}$$

between the three types of payoffs, the interest rate paid, and the share of cash flows) the “mix” of debt and equity features can be varied, making the silent partnership rather more like equity—or rather more like debt.

A recent study of German venture capitalists (by Hartmann-Wendels, Keienurg, and Sievers) reveals that the forms of funding do seem to be chosen to address the moral hazard and adverse selection issues.

1. For start-up and high technology firms where the probability of failure is high and the moral hazard and adverse selection problems are severe, these firms are financed largely by equity or a combination of equity and silent partnerships with more equity-like features. If you look back to issues A and B at the beginning of this Strategy Session, you will see that this is exactly what the theory predicts.
2. Where the venture capitalist is very experienced, he is able to cherry pick the good firms and, having some confidence in their future prospects, he will wish to share the upside with equity, and equity-like, financing. And indeed, this is what is found in the data.
3. They also found that when the economy was in recession, investors became more conservative and

more downside protection was sought by investors. Thus, more use of debt, and debt-like, financing was used. As the economy improved, investors gravitated back to more equity financing.

¹Convertible debt is debt, that can be converted into equity if the investor chooses to do so. Investors will make this conversion if the firm does well. But if the firm is not performing well, investors will not exercise their option to make a conversion, and the instrument will continue to be debt.

²With limited liability, the shareholders can “walk away” from the firm should it become bankrupt. In other words, the value of the shares will not be negative. When the firm is bankrupt, the creditors are not paid in full; but they do receive any remaining value left in the firm. For example, if the firm is worth \$50 million but owes \$100 million, then the shareholders will receive nothing and the creditors will receive the whole \$50 million; or “50 cents on the dollar.” Since the shareholders can “walk away,” then the creditors must bear the transaction costs of the bankruptcy.

Sources: Thomas Hartmann-Wendels, Georg Keienburg, and Soenke Sievers, 2011, “Adverse Selection, Investor Experience and Security in Venture Capital Finance: Evidence from Germany,” *European Financial Management*, Vol. 17, pp. 464–499; Douglas J. Cumming 2006, “Adverse Selection and Capital Structure: Evidence from Venture Capital.” *Entrepreneurship Theory and Practice*, Vol. 30, pp. 155–184. Available at SSRN: <http://ssrn.com/abstract=261693> or doi:10.2139/ssrn.261693.

For the low-risk drivers, the expected utilities are calculated using the low-risk probabilities

$$\text{No insurance utility} = (0.75)(125)^{0.5} + (0.25)(125 - 100)^{0.5} = 9.635$$

$$\text{Policy 1 utility} = (125 - 75)^{0.5} = 7.071$$

$$\begin{aligned}\text{Policy 2 utility} &= (0.75)(125 - 2.5)^{0.5} + (0.25)(125 - 100 - 2.5 + 10)^{0.5} \\ &= 9.726\end{aligned}$$

Note that the two types of drivers choose different policies. For high-risk drivers the best choice is policy 1, which offers them a utility of 7.071 (versus 7.043 for policy 2 and 6.545 for buying no insurance). But for low-risk drivers policy 2 is the best choice, offering them a utility of 9.726 (versus 7.071 for policy 1 and 9.635 for no insurance). In making these choices, individual policyholders act in their self-interest (maximizing expected utility) and use their private information about their loss probabilities. Note also that this self-selection has taken place without managers initially being able to identify which policyholder is at low risk and which is at high risk of accidents. Hence we can call this a **self-selection menu**.

Self-selection menu When buyers act in their own self-interest and use their private information about their loss probabilities to select policies.

There are other interesting features about this solution chosen by managers. When we introduced this adverse selection problem earlier in the chapter, only one policy survived: policy 1. In that earlier analysis, the good drivers preferred self-insurance to buying policy 1. Now with the self-selection menu of policies 1 and 2, managers make the good drivers better off (they prefer the newly introduced policy 2 to self-insurance) while bad drivers are in the same position (they continue to buy policy 1). This is a clear improvement; managers improved the lot of good drivers at no cost to the bad drivers. So this menu has at least partly salvaged the marketplace by enabling low-risk drivers to get insurance.

Finally, managers need to know whether such a menu is feasible. The most immediate worry is whether managers will want to offer both policies. Because the policyholders self-select, the insurer breaks even with each policy:

- Only high-risk drivers buy policy 1. The expected claims are 0.75 (100), which matches the premium of 75.
- Only low-risk drivers buy policy 2. Hence expected claims are 0.25 (10), which matches the premium of 2.5.

Therefore, a competitive insurer can offer this choice and survive.

This solution to the adverse selection problem is called a **separating equilibrium**. It is in the interests of the uninformed insurer to offer such a choice; and the choice induces self-selection by the policyholders into their respective risk types. But note that information asymmetry still imposes a cost. The insurance coverage in policy 2 does not offer complete protection to the low-risk drivers. Although it is clearly better than no insurance, the coverage is rationed.

Separating equilibrium This solution to adverse selection induces policyholders to select their respective risk types.

USING EDUCATION AS A SIGNAL: ADVERSE SELECTION IN THE JOB MARKET

An early, and somewhat surprising, application of looking at the incentives of information was presented by the economist Michael Spence (who, as we mentioned, shared the 2001 Nobel Prize). The problem he had in mind is straightforward enough; the solution is quite clever. How can managers predict which job applicants will have good work skills? We show that applicants use education to send to potential employers a signal of labor market skills that cannot be easily measured by employers. Of course schools and universities also teach something.

Applicants have some self-awareness about their job market skills. They know their ambitions, how organized they are, and whether they are prepared to work hard and long hours. They also know something about their people skills and intellectual ability.

These traits are initially not known to a potential employer, who may learn about employee abilities only gradually and over a long period. So there is an information asymmetry. As in the insurance and annuity markets, simply asking applicants to reveal their private information does not work. Most applicants know how to prepare a résumé and hone their interview techniques. So managers need a credible method for separating those who know they have good job skills from those who lack them.

If managers believe there is a relationship between job skills and academic performance, they can craft a way to find those with good work skills. This need not be a perfect relationship; but on average, people with good job skills have an easier time overcoming academic hurdles. Consider an undergraduate degree that requires students to pass 30 courses. Many students do this in four years, but some take three years and others may take five or six because they have to repeat courses or have a smaller course load. The cost of getting a degree, therefore, varies according to the length of time it takes. These costs include the direct costs of paying for education and the opportunity costs of losing wages while not working. Direct costs mount as courses are repeated, and lost wages climb as students take longer to finish.

Consider the following example. Average direct and indirect costs per course are

High-quality job skills = \$2,000 per course

Low-quality job skills = \$3,450 per course (they take longer to finish)

Now, assume that if employers knew the skill levels of applicants, they would pay the following wages

High-quality job skills = \$50,000 per year for 5 years

Low-quality job skills = \$30,000 per year for 5 years

After five years, the employer can figure out an employee's job skills.

Of course employers do not know each person's job skills; but suppose employers make the following offer: All those who have taken at least x courses are paid a

salary of \$50,000, and those that took fewer than x courses are paid \$30,000. Managers need to calculate a value of x to separate the high- and low-skilled workers.

Students know this offer stands, and they have to decide how many courses to take. So they rationally look at the costs and benefits of different degree programs. The benefit of completing x courses to any student, of high or low skill, is a wage increase from \$30,000 to \$50,000 for five years, giving a total benefit of \$100,000 (we ignore the time value of money here). But the costs differ, being \$2,000 per course and \$3,450 per course for the low- and high-skilled people.

High skill:

Benefit of achieving x courses = \$100,000

Cost of achieving x courses = \$2,000 times x

The benefit exceeds the costs if x is less than 50.

Low skill:

Benefit of achieving x courses = \$100,000

Cost of achieving x courses = \$3,450 times x

The benefit exceeds the costs if x is less than 29.

So the employers choose a level of x between 29 and 50. Say they choose $x = 30$, which is the typical four-year degree program. People with low-quality skills do not choose to take the 30-course degree (the cost is $\$3,450 \times 30 = \$103,500$, and the benefit is only \$100,000). But high-skilled students take the 30-course degree (the cost is $\$2,000 \times 30 = \$60,000$, and the benefit is \$100,000).

By cleverly setting the 30-course standard, managers persuade people to reveal their hidden information as they choose their education. Universities now provide two functions. In addition to actual teaching, they screen people according to employment skills and endorse these skills to potential employers. Another way of thinking about this is provided by our MBA students. When asked why they go for the Wharton MBA, many reply, “Employers pay more for a Wharton MBA because they know that anyone accepted to this program and obtaining the degree must be good.” Therefore, if a university wishes to signal to the market that its graduates are of high quality, it must set its standards sufficiently high to discourage low-skilled people.

Of course, education has other functions in addition to sorting people according to job skills. Ideally people learn something as well.⁵

USING WARRANTIES AS SIGNALS: ADVERSE SELECTION IN THE PRODUCT MARKET

Though a consumer selects a product (and pays for it) at the time of purchase, there are many products whose quality cannot be determined until they are con-

5. All the authors of this book are professors, and their jobs might depend on this caveat.

sumed. These products are called *experience goods*. Examples of such goods include autos, appliances, and consumer electronics. As we saw in our used car example, since consumers cannot determine quality until after their purchases, there is an incentive for low-quality producers to advertise their products as high-quality. If consumers cannot determine the true quality of the good prior to purchase, they are unwilling to pay the price of a high-quality good.

In such situations, there is an incentive for managers of firms that produce high-quality goods to signal their true quality. Hence, they want to take actions to create a separating equilibrium so consumers can accurately determine product quality. One of the most common methods managers use to create this separation is the use of a product warranty. As we see below, if constructed correctly, a warranty is a credible signal of product quality. And, there is empirical evidence showing that consumers do recognize this and are willing to pay a higher price for products that carry a credible warranty.

How Managers Can Construct Warranties to Mitigate Adverse Selection

Imagine that managers of rival firms introduce a new product. One product is truly of high-quality, while the other is of low-quality. Consumers have a reservation price for a high-quality product and a different (and lower) one for the low-quality product, if they could determine quality prior to purchase. It costs more to produce the high-quality product and less to produce the low-quality product.

If consumers could credibly determine product quality, we'd expect a separate market for high-quality products and a separate market for low-quality products. Managers of the firm producing the high-quality product want to construct a separating mechanism so consumers can determine product quality prior to purchase. Here is how they can do so using a product warranty.

In the following analysis, both firms might offer a warranty, but they would be of different lengths. If both firms offer a warranty, consumers believe that whichever product has the longer warranty is the higher-quality product.

Consumers' reservation price for the high-quality good is P_H and their reservation price for the low-quality good is P_L (where $P_H > P_L$). It costs the producer of the high-quality product a constant C_H to produce a unit and it costs the producer of the low-quality product a constant C_L to produce a unit (where $C_H > C_L$). Naturally, high-quality goods are more expensive to produce. It costs the producer of the high-quality product W_H /year to honor a full-coverage warranty and it costs the producer of the low-quality product W_L /year to honor a full coverage warranty (where $W_L > W_H$, because the low-quality item needs repair/replacement more often). Y is the number of years of the warranty. Producers wish to maximize their per unit profit.

Now recall that consumers view the longest warranty as a signal of high-quality. The longest warranty a high-quality producer can afford to give is one

PROBLEM SOLVED: Determining Warranty Length

Tole Brothers is a new entrant into the house-building market. They haven't been in the market long enough to gain a reputation for the quality of the homes they build. They think of themselves as builders of high-quality homes.

In the area where they're building, a high-quality home will sell for \$500,000. It will cost a builder \$250,000 to build a high-quality house. A rival firm, Quality Builders can construct a house that looks like a high-quality house but soon after occupation, the owners will begin to notice the defects. If the buyers can identify such a home, they would pay \$400,000 for it. It would cost Quality Builders \$200,000 to construct such an inferior home.

Managers at Tole Brothers believe they can signal their homes are high-quality by issuing a warranty on the home (at no cost to the buyer) that protects the home for W years against any defects. If a defect is found, they will repair it for free. Managers estimate the expected cost of this warranty to Tole

Brothers is \$10,000 per year, per home. If managers at Quality Builders want to offer an analogous warranty, it will cost them \$25,000 per year, per home, because the likelihood of defects is much greater in their poorly constructed homes. Should managers at Tole Brothers offer a warranty? If yes, for what duration?

If managers at Tole Brothers choose not to offer a warranty, their houses are perceived as inferior and will sell for \$400,000. Since they cost \$250,000 to build, Tole Brothers will make \$400,000 – 250,000 or \$150,000 per house.

If managers at Tole Brothers choose to offer a warranty, their houses are perceived as high-quality and sell for \$500,000. It costs them \$250,000 to build the house and a warranty costs \$10,000 per year, per house to service. Thus, a warranty of W years would cost $10,000W$. Thus, Tole Brothers' profits with a warranty is \$500,000 – \$250,000 – \$10,000 W = \$250,000 – \$10,000 W .

where the profit from signaling high-quality just equals the profit from not signaling at all, that is, low-quality.⁶ This will have length Y_H where

$$P_H - C_H - W_H Y_H = P_L - C_H$$

The longest warranty a low-quality producer can afford to give is one where the profit from signaling high-quality just equals the profit from not signaling at all, that is, low-quality. This will have length Y_L where

$$P_H - C_L - W_L Y_L = P_L - C_L$$

Since $W_L > W_H$, then $1/W_H > 1/W_L$, and so $Y_H > Y_L$.

Thus, the high-quality producer can afford to out-warranty the low-quality producer. But now, think strategically. If the high-quality producer issued a warranty of Y_H years, managers would dissipate all the gains of being recognized as high-quality, that is, they would earn the same profit as if they had been perceived to be low-quality. At the same time, the low-quality producer can't afford to offer

6. If a producer is unable to offer the longest warranty, consumers will view the product to be of low quality. Thus, if the longest warranty is not offered, it is not worth the cost of issuing a shorter warranty (the shorter warranty would be insufficient to persuade consumers of quality).

By setting Tole Brothers' profits with and without a warranty equal and solving for W , we get

$$\begin{aligned} \$250,000 - \$10,000W &= \$150,000 \\ \text{or } \$10,000W &= \$100,000 \text{ or } W = 10. \end{aligned}$$

Thus, managers at Tole Brothers could offer a warranty of up to 10 years and make more money per house than if they offered no warranty. But do they have to?

If managers at Quality Builders do not issue a warranty, their houses are perceived as inferior and will sell for \$400,000. Since they cost \$200,000 to build, Quality will earn \$400,000 - \$200,000 or \$200,000 per house.

If managers at Quality decide to offer a warranty, their houses are perceived as high-quality and will sell for \$500,000. It costs them \$200,000 to build the house and the warranty costs \$25,000 per year, per house to service. Thus, a warranty of W years will cost Quality \$25,000 W . Hence the managers earn a profit of \$500,000 - \$200,000 - \$25,000 W or \$300,000 - \$25,000 W .

By setting Quality's profits with and without a warranty equal and solving for W , we get

$$\begin{aligned} 300,000 - 25,000W &= 200,000 \\ \text{or } 25,000W &= 100,000 \text{ or } W = 4 \end{aligned}$$

Thus, managers at Quality Builders could offer a warranty of up to 4 years and make more money per house than if they offered no warranty.

Managers at Tole Brothers now must set their strategy. Suppose they offer a warranty of slightly more than 4 years. Managers at Quality then will not offer any warranty (since they make more money with no warranty). Quality will not offer a warranty with their houses and managers will earn \$200,000 per house.

Managers at Tole Brothers will earn a profit of \$500,000 - \$250,000 - \$10,000(4+) = \$250,000 - \$40,000+ or just under \$210,000 per house. This profit exceeds the \$150,000 they'd make if they didn't issue a warranty and signal low quality.

a warranty of more than Y_L years (because if managers issued a warranty of even Y_L years, they would dissipate all the gains of being perceived as high-quality). So all the high-quality producer has to do is just out-warranty the low-quality producer. A warranty of $Y_H = Y_L + \varepsilon$ will do the trick. The high-quality producer can afford to give it and the low-quality producer cannot afford to give it and so the proposed length of warranty is credible. The low-quality producer will not issue a warranty (since it doesn't provide a positive signal and managers would incur costs of honoring the warranty). Notice the similarity to auctions. The person, who values the item most, needs just to outbid the person who values the item second most.

Does this help explain why buyers are willing to pay an auto dealer a higher price for a pre-owned certified auto with a warranty than for an apparently similar auto they purchase from a private seller? Or, why consumers pay a higher price for electronic products from a store that offers a warranty relative to apparently similar items on eBay that lack such a warranty?

PROBLEM SOLVED: How a Venture Capitalist Can Address Moral Hazard and Adverse Selection Problems

In general, research has shown that firms with different earnings characteristics will favor different types of financing to pay for their investment projects. Firms with low expected earnings will tend to favor equity financing; firms with highly volatile earnings will tend to favor debt; and firms with low earnings volatility will tend to favor convertible securities. If investors do not know the firm's earnings features, there will be an adverse selection issue. We will close this chapter by looking at a more complex issue that combines both adverse selection and moral hazard. We consider entrepreneurs who seek financing from venture capitalists (investors). However, investors do not have good information about the quality of the entrepreneurs (adverse selection) and entrepreneurs can make investment choices after they have secured funds to pay for them (moral hazard).

Suppose there are two types of risk-averse entrepreneurs; we will call them HIGH and LOW reflecting differences in their respective entrepreneurial abilities. Each type can start up a firm which will operate for one period. At the end of the period, the value is determined and this value is divided between the investor who provided financing and the entrepreneur who started the firm.

HIGH has a choice between investing in a safe project, with a certain cash value of \$260,000, or a risky project that has the following cash value at the end of the year

- 0.60 chance of \$100,000
- 0.40 chance of \$500,000

LOW has a choice between investing in a safe project, with a certain cash value of \$260,000, or a risky project that has the following cash value at the end of the year

- 0.75 chance of \$100,000
- 0.25 chance of \$500,000

All cash values are realized in one year. The difference is, of course, that the HIGH quality entrepreneur has a higher chance of being successful.

To undertake any of these projects requires an up-front investment of \$200,000. Assume that the entrepreneur has no money and must seek this from an investor. The funding can come as either debt or equity.

- **DEBT:** With debt funding, the full \$200,000 is borrowed. Debt should be repaid at the end of the period. This will involve payment of principal (\$200,000) and interest at 22.5% (\$45,000) requiring a total repayment of \$245,000. We will see presently why the interest rate is so high.
- **EQUITY:** The investor provides the required funding of \$200,000 in exchange for an 85% share in the final value.

Entrepreneurs are risk averse with a utility function of $U = \sqrt{W}$ where W is their share of the firm's terminal value (after repayment of principal and interest—or the 85% dividend—depending on which form of finance is chosen). However, there is limited liability, and if the firm does borrow and ends up with a value of less than \$245,000, it will be bankrupt. In this case, the entrepreneur gets nothing, and the investor is repaid the full cash flow, whatever that happens to be.

Investors are risk neutral. While investors know that there are some entrepreneurs of HIGH and some of LOW quality, they are unable to tell which is which. Nor does the investor know which project the entrepreneur will choose; the risky one or the safe one. We assume that the entrepreneur is better informed and does know whether he is of the HIGH or LOW type.

The Entrepreneurs' Choices

1. Which project will the HIGH entrepreneur choose (the safe or risky one) if he funds with debt?
2. Which project will the HIGH entrepreneur choose (the safe or risky one) if he funds with equity?
3. Which project will the LOW entrepreneur choose (the safe or risky one) if he funds with debt?
4. Which project will the LOW entrepreneur choose (the safe or risky one) if he funds with equity?

QUESTION 1 Consider HIGH's choice if he funds with debt

With the SAFE project, he nets the firm's value of \$260,000 minus the repayment of principal and interest which is \$245,000. Thus, his expected utility is

$$EU(\text{HIGH}; \text{SAFE}) = \sqrt{260,000 - 245,000} = 122.5$$

With the RISKY project, he nets the firm's value of either \$100,000 or \$500,000 minus the repayment of principal and interest which is \$245,000. However, with limited liability, the entrepreneur cannot be forced to repay the debt if the firm is bankrupt, thus he will end up with ZERO. Thus, his expected utility is

$$\begin{aligned} EU(\text{HIGH}; \text{RISKY}) &= 0.6\sqrt{\text{ZERO}} + 0.4\sqrt{500,000 - 245,000} \\ &= 202.0 \end{aligned}$$

HIGH will therefore choose the RISKY project if offered debt funding.

QUESTION 2 Consider HIGH's choice if he funds with equity

With the SAFE project, he nets 15% of the firm's value of \$260,000. His expected utility is

$$EU(\text{HIGH}; \text{SAFE}) = \sqrt{260,000 \text{ times } 0.15} = 197.5$$

With the RISKY project, he nets 15% of either \$100,000 or \$500,000. His expected utility is

$$\begin{aligned} EU(\text{HIGH}; \text{RISKY}) &= 0.6\sqrt{100,000 \text{ times } 0.15} \\ &\quad + 0.4\sqrt{500,000 \text{ times } 0.15} \\ &= 183.0 \end{aligned}$$

HIGH will therefore choose the SAFE project if offered equity funding.

QUESTION 3 Consider LOW's choice if he funds with debt

With the SAFE project, he nets the firm's value of \$260,000 minus the repayment of principal and interest which is \$245,000. His expected utility is

$$EU(\text{LOW}; \text{SAFE}) = \sqrt{260,000 - 245,000} = 122.5$$

With the RISKY project, he nets the firm's value of either \$100,000 or \$500,000 minus the repayment of principal and interest which is \$245,000. However, with limited liability, the entrepreneur cannot be forced to repay the debt if the firm is bankrupt, thus he will end up with ZERO. His expected utility is

$$\begin{aligned} EU(\text{LOW}; \text{RISKY}) &= 0.75\sqrt{\text{ZERO}} \\ &\quad + 0.25\sqrt{500,000 - 245,000} = 126.2 \end{aligned}$$

LOW will therefore choose the RISKY project if offered debt funding.

QUESTION 4 Consider LOW's choice if he funds with equity

With the SAFE project, he nets 15% of the firm's value of \$260,000. His expected utility is

$$EU(\text{LOW}; \text{SAFE}) = \sqrt{260,000 \text{ times } 0.15} = 197.5$$

(Continued)

PROBLEM SOLVED: How a Venture Capitalist Can Address Moral Hazard and Adverse Selection Problems (*Continued*)

With the RISKY project, he nets 15% of either 100,000 or 500,000. His expected utility is

$$\begin{aligned} EU(\text{LOW}; \text{RISKY}) &= 0.75\sqrt{100,000 \text{ times } 0.15} \\ &\quad + 0.25\sqrt{500,000 \text{ times } 0.15} \\ &= 160.3 \end{aligned}$$

LOW will therefore choose the SAFE project if offered equity funding.

You may note several things.

1. There is a moral hazard problem. Because of limited liability, the entrepreneur has the chance to walk away from any debt, rather than repaying in full. Because of this, entrepreneurs who use debt financing tend to favor more risky projects—because funding with debt allows them to keep all the upside profit if things go well—but to leave the debt largely unpaid if things turn out badly. The effect of this perverse incentive is to make the entrepreneur's choice of project dependent on the form of financing. Both would prefer the safe project if offered equity financing, but would switch to the risky project if funds were obtained with debt. The opportunity to borrow and then walk away if things go wrong is very seductive and persuades entrepreneurs to go for the risky projects and let the creditors take the hit.
2. There is an adverse selection problem. The investor doesn't know the entrepreneur's type, but the entrepreneur does indeed know who he is. But note that the LOW types would prefer to have equity financing—in which case they would choose the safe

project and end up with utility of 197.5. In contrast, the HIGH types would prefer to have debt financing and they would choose the risky project, which would give them utility of 202. Thus, if the investor offered a free choice of the type of funding, then the different types would reveal themselves in their choices. But as we will see, it may be dangerous for the investor to make such an offer.

The Investor's Decisions

Now consider the investor. The investor does not know whether she is lending to a HIGH or LOW type of entrepreneur. Suppose one day, an entrepreneur (of unknown type) approaches her for funding. Which type of financing should the investor offer (if any)?

If she offers EQUITY financing then

HIGH will choose the SAFE project. The investor will receive 85% of \$260,000. The investor has an expected return of

$$\begin{aligned} 0.85 \text{ times } 260,000 &= \$221,000 \\ &\quad (\text{minus the } \$200,000 \text{ invested}) \end{aligned}$$

LOW will also choose the SAFE project. Thus the investor's return is

$$\begin{aligned} 0.85 \text{ times } 260,000 &= \$221,000 \\ &\quad (\text{minus the } \$200,000 \text{ invested}) \end{aligned}$$

If she offers DEBT financing then

HIGH will choose the RISKY project. The investor will receive either the full repayment of interest

and principal (\$245,000) if the firm is solvent, or the full remaining value of the firm (\$100,000) if the firm is bankrupt. Then the investor will have an expected return of

$$0.6 (100,000) + 0.4 (245,000) = 158,000$$

(minus the \$200,000 invested)

LOW will also choose the RISKY project. Thus, the investor's return is

$$0.75 (100,000) + 0.25 (245,000) = 136,250$$

(minus the \$200,000 invested)

Thus with both types, the investor will lose money by offering the debt financing. The investor should therefore confine itself to offering equity funding. This disciplines both firms to choose safe projects and thus resolves the moral hazard problem. It also renders the adverse selection problem irrelevant, the difference in types only concerned with their ability to manage a risky project, and now they choose not to do so.

Additional Thoughts

1. Would debt financing ever be relevant? Yes, it may be if the risk that the entrepreneur would choose to take with debt financing was modest and the rewards great. Suppose there were a third type of investor called BLUE CHIP. This investor can also either undertake the safe or risky projects. But the difference is that if BLUE CHIP chooses the risky investment, he has a 75% chance of succeeding. You

should work through this example (along the lines of the previous calculations) and you will find that

- BLUE CHIP would choose the risky project with either debt or equity financing (if the funding terms were the same as before; 22.5% interest for debt and an 85% dividend for equity)
- BLUE CHIP would nevertheless prefer debt to equity
- The investor would still be able to make money by lending to BLUE CHIP at the nominal 22.5% interest. The investor would in fact prefer an equity stake of 85% to lending. However, BLUE CHIP may be able to raise equity funding more cheaply, i.e., by offering somewhat less than an 85% share of the firm.

2. In this example, we have offered only two types of financing, debt and equity. You might also contemplate intermediate types such as convertibles. There are various types of convertibles that can morph from debt to equity, or vice versa; and there are preferred shares. We will not get into this here except to note that these allow greater flexibility in how the upside and downside risk are divided between the investor and the entrepreneur; and thus allow fine tuning in the resolution of the moral hazard and adverse selection problems.

3. Also in this example, we have fixed the terms of the debt and equity issues. In practice, competition would force these returns to (wards) some equilibrium levels *given the firms that actually use them*. This, of course, adds complexity to the solution.

SUMMARY

1. Adverse selection arises when one party to a contractual or economic relationship knows more than the other. For example, a seller of a used car knows more than a buyer, a policyholder knows more than an insurance company, and a borrower knows more than a lender. This puts the uninformed party at a disadvantage. For example, managers at insurance companies are unable to distinguish between safe and unsafe drivers; therefore premiums are averaged over both types. This means that good drivers subsidize bad ones, and many good drivers may be tempted to cancel their insurance to avoid this subsidy. In the extreme case, adverse selection may bring a market crashing down as all the low-risk people are priced out of the market. Similarly, all sellers of high-quality used cars might decide to keep them rather than sell at a price reflecting average quality.

2. There are tactics managers can use to mitigate the effects of adverse selection. The obvious one is to become informed. We saw that while adverse selection exists in the annuity market, managers at life insurance companies seem to have been successful in removing the information problem by medical examinations.

3. Managers can also design a menu of contracts to let others reveal their asymmetric information; the contracts in the menu have differential appeal to different customers. For example, in insurance, a high-priced policy might offer full insurance; a low-priced policy might cover only part of the damage. Bad drivers worry about the partial coverage because they know they are likely to have an accident. On the other hand, good drivers might like the partial coverage because it is cheap, and they figure they are unlikely to be in a crash. In labor markets education is used as a signal of worker productivity; and producers might use warranties to distinguish their products from those of rivals.

PROBLEMS

1. Sellers of used cars know the cars' quality, but buyers do not. Imagine that used Toyota Corollas are worth \$10,000 if they are of high quality and \$5,000 if they are of poor quality. Although buyers may not know the quality of a specific car, they do know that 25% will be of poor quality. In such a market, what cars will be sold on the secondhand market and at what price?
2. The market for digital cameras is relatively new. Ajax Inc. produces what it regards as a high-quality digital camera. Knockoff Inc. produces what it regards as a low-quality digital camera. However, because the market is so new, reputations for quality have not yet developed, and consumers cannot tell the quality difference between an Ajax digital and a Knockoff digital just by looking at them.

If consumers knew the difference, they'd be willing to pay \$200 for a high-quality camera, and they'd be willing to pay \$100 for a low-quality camera. It costs Ajax \$85 to produce a high-quality camera, and it costs Knockoff \$55 to produce a low-quality camera.

A recent MBA hire at Ajax suggests that Ajax could differentiate its camera from Knockoff's by offering a full-coverage warranty (which would fully cover any defect in the camera at no cost to the customer). The MBA estimates that it would cost Ajax \$20 per year to offer such a warranty. The MBA also estimates that it would cost Knockoff \$40 per year should Knockoff attempt to copy Ajax's warranty strategy. Consumers will feel that the camera with the longest warranty is high-quality and that with the shortest warranty is low-quality. The camera companies want to maximize the profit per camera.

What is Ajax's profit per camera in the digital camera market?

3. No-State Insurance Company has made the following estimate of auto damage for several groups of potential customers who own cars worth \$10,000. There are an equal number of customers in each group. No-State is risk-neutral.

Group	Initial Value of Car	Probability of Accident that Devalues Car to \$5,000
A	\$10,000	0.2
B	\$10,000	0.3
C	\$10,000	0.4

State regulation mandates that every customer must pay the same premium regardless of his or her group, and this premium must be sufficient to cover all expected claims from those who purchase insurance from No-State. There are no additional costs to the company other than paying off claims.

All consumers have the following utility function (U)

$$U = W^{0.5}$$

W is the consumer's wealth as represented by the value of the car.

What premium should No-State offer for full-coverage insurance?

4. Some people are good drivers, and others are bad drivers. The former have a 10% chance of crashing their cars, and the latter have a 30% chance. All have a total wealth of 400, but this will fall to 100 if they crash their cars. In other words, each will lose 300 of wealth if they crash. You are an insurance company manager who wishes to offer a pair of policies to all drivers. Each policy is designed to break even (zero profit) given the people who choose to buy that policy. The first policy has a premium of 90 and covers all losses (it will pay 300 in the event of a crash). The second policy has a premium of 5 and will

pay 50 in the event of a crash. Who will buy which policy? Will the insurance company make a profit, break even, or lose money?

Each person has a utility function as follows

$$\text{Utility} = (\text{Wealth})^{0.5}$$

5. Consider a market for annuities for 70-year-old men in which people differ in terms of both their expected remaining years of life and their risk preferences. Of the population of 200, half have a life expectancy of 9 years and the remaining half have a life expectancy of 11 years. We can express risk preference in the following way. The risk people are worried about is that of running out of wealth before they die. The more risk-averse you are, the higher the up-front price you are willing to pay for the annuity. More risk-averse people are willing to pay 1.3 times x times A , where x is the expected years of life remaining and A is the dollar amount paid each year to the annuitant. Less risk-averse people are willing to pay only 1.1 times x times A . Assume that of the 100 people in each health group, half are more risk-averse and half less risk-averse.

The annuity firm sells an annuity of \$50,000 per year for as long as the buyer lives, and the price of the annuity is \$550,000. Because the annuity firm cannot tell whether any applicant has a short or long life expectancy, it must accept any application for its product. What is the expected profit of the annuity firm? (You may ignore discounting in this example.)

EXCEL EXERCISE: ADVERSE SELECTION

Adverse selection occurs when there is hidden information. The five individuals below each have a different probability of a catastrophic health event. The difficulty exists because while every individual knows his or her probability, the health insurer does not (the hidden information).

There are five types of individuals (A, B, C, D, and E). For simplicity, assume there is one of each type. Each has initial wealth (W) of 10. They are all risk neutral with a utility (U) function of $(U) = W$. Individual A has a 10% chance of a catastrophic event that takes his wealth to 0. Individual B faces the same catastrophic event, but with probability at 20%. Individual C faces the catastrophic event with probability at 30%, D with 40%, and E with 50%.

An insurance company could offer the individuals full-coverage insurance. The industry is regulated, and the regulatory agency requires that the same price (premium) be charged to each individual and that the premium be actuarially fair, that is, result in the sum of all premium income equaling the sum of all expected payouts.

Under these conditions, what is the expected premium for insurance in this market, and who buys insurance at that premium?

Enter 10 in cell A1 and then click on the lower right-hand corner and drag down to cell A5. This is each individual's initial wealth.

Enter 0.1 in cell B1, 0.2 in cell B2, 0.3 in cell B3, 0.4 in cell B4, and 0.5 in cell B5; Column B is the probability that each individual faces the catastrophic event.

Enter $=A1*B1$ in cell C1. Then click and drag until cell C5. Column C represents the expected claims of each individual. Enter $=\text{Sum}(C1:C5)$ in cell C6. This is the total claims to be submitted by the five individuals. Enter $=C6/5$ in cell C7. This is the actuarially fair insurance rate to be charged in the market if all five individuals buy insurance.

Enter $=A1-C7$ in cell D1, $=A2-C7$ in cell D2, $=A3-C7$ in cell D3, $=A4-C7$ in cell D4, and $=A5-C7$ in cell D5. Column D shows the net income of each individual if they buy the full coverage insurance at the premium of C7.

Enter $=A1-C1$ in cell E1. Then click and drag until cell E5. Column E shows the net income of each individual if they self-insure.

Then compare the respective entries in columns D and E, that is, D1 versus E1, etc. If the entry in column E exceeds the entry in column D, the individual will self-insure. If the opposite occurs, the individual will purchase the insurance. If the numbers are equal, we'll assume that they will purchase the insurance. In this case, you will see that A and B will self-insure. This leaves C, D, and E in the potential insurance market.

Enter $=C3$ in cell F3, $=C4$ in cell F4, and $=C5$ in cell F5. Then enter $=\text{Sum}(F3:F5)$ in cell F6. F6 represents the total of the claims to be submitted if C, D, and E are insured. Enter $=F6/3$ in cell F7. F7 represents the actuarially fair insurance rate if C, D, and E are insured.

Enter $=A3-F7$ in cell G3, $=A4-F7$ in cell G4, and $=A5-F7$ in cell G5. Column G represents the net income of the individual if he or she purchases insurance at the actuarially fair rate.

Enter $=A3-C3$ in cell H3, $=A4-F4$ in cell H4, and $=A5-F5$ in cell H5. Column H represents the net income to each individual if he or she self-insures.

Then compare the respective entries in columns G and H, that is, G3 versus H3, and so on. If the entry in column H exceeds the entry in column G, the individual will self-insure. If the opposite occurs, the individual will purchase the insurance. If the numbers are equal, we'll assume that the individual will purchase the insurance. In this case, you will see that C will self-insure. This leaves D and E in the potential insurance market.

Enter $=C4$ in cell I4 and $=C5$ in cell I5. Then enter $=\text{Sum}(I4:I5)$ in cell I6. I6 represents the total of the claims to be submitted if D and E are insured. Enter $=I6/2$ in cell I7. I7 represents the actuarially fair insurance rate if D and E are insured.

Enter $=A4-I7$ in cell J4 and $=A5-I7$ in cell J5. Column J represents the net income of the individual if he or she purchases insurance at the actuarially fair rate.

Enter $=A4-I4$ in cell K4 and $=A5-I5$ in cell K5. Column K represents the net income to each individual if he or she self-insures.

Then compare the respective entries in columns J and K, that is, J4 versus K4, and so on. If the entry in column K exceeds the entry in column J, the individual will self-insure. If the opposite occurs, the individual will purchase the insurance. If the numbers are equal, we'll assume that the individual will purchase the insurance. In this case, you will see that D will self-insure. This leaves only E in the potential insurance market.

The potential claims from individual E are shown in C5. Thus, the actuarially fair rate is C5. Individual E will pay that rate because his or her net income is the same whether he or she insures or self-insures. Thus, the market has an insurance premium of C5, with only individual E purchasing insurance (or, if we don't assume that ties in net income would purchase insurance, flipping an unbiased coin to determine whether or not to purchase).

This is the "lemons" problem in insurance, in which the "bad" customers drive the "good" customers out of the market.

Suppose that our individuals were as above but with $U = W^{0.5}$; that is, they were risk-averse. How would this situation play out?

Enter 10 in cell A9 and click and drag until cell A13.

Enter 0.1 in cell B9, 0.2 in cell B10, 0.3 in cell B11, 0.4 in cell B12, and 0.5 in cell B13.

Enter $=A9*B9$ in cell C9 and click and drag until cell C13. Column C is the expected claims of each individual. Enter $=\text{Sum}(C9:C13)$ in cell C14. This is the expected total claims submitted to the insurance company if all individuals purchase insurance. Enter $=C14/5$ in cell C15. This is the actuarially fair insurance rate if each individual buys insurance.

Enter $=(A9-C15)^{0.5}$ in cell D9, $=(A10-C15)^{0.5}$ in cell D10, $=(A11-C15)^{0.5}$ in cell D11, $=(A12-C15)^{0.5}$ in cell D12, and $=(A13-C15)^{0.5}$ in cell D13. Column D is the expected utility of each individual if he or she purchases insurance at the actuarially fair rate.

Enter $=(1-B9)*A9^{0.5}+B9*0^{0.5}$ in cell E9 and click and drag until cell E13; Column E is the expected utility of each individual if he or she self-insures.

Compare the respective entries of columns E versus D, that is, E9 versus D9. If the E column entry exceeds the respective D column entry, the individual will self-insure (because the expected utility from self-insuring exceeds the expected utility from buying insurance). In this case, this is true for individual A. But for individuals B through E, the column D entry exceeds the respective column E entry, so they will wish to buy insurance.

Enter 10 in cell F10 and click and drag down to cell F13.

Enter 0.2 in cell G10, 0.3 in cell G11, 0.4 in cell G12, and 0.5 in cell G13.

Enter $=F10*G10$ in cell H10 and click and drag until cell H13. Column H represents the expected claim filed by each individual. Enter $=\text{Sum}(H10:H13)$

in cell H14. H14 is the total sum of claims if individuals B, C, D, and E purchase insurance. Enter $=H14/4$ in cell H15. H15 is the actuarially fair insurance premium if individuals B, C, D, and E purchase insurance.

Enter $=(F10-H15)^{0.5}$ in cell I10, $=(F11-H15)^{0.5}$ in cell I11, $=(F12-H15)^{0.5}$ in cell I12, and $=(F13-H15)^{0.5}$ in cell I13. Column I is the expected utility of each individual if he or she buys insurance at the actuarially fair rate.

Enter $=(1-G10)*F10^{0.5}+G10*0^{0.5}$ in cell J10 and click and drag until cell J13. Column J shows the expected utility of each individual if he or she self-insures.

Compare the respective entries in column J and column I. If the column I entry exceeds the respective column J entry, then that individual buys insurance. That is true in this case for each of individuals B through E. Thus, risk-aversity can preserve a significant share of the insurance market from adverse selection. The insurance market has four of the five individuals participating at an actuarially fair premium of H15.

If the utility function of the individuals had been $U = W^{0.25}$, all individuals would have purchased insurance at the actuarially fair price of C15. You can try this yourself. Use the cells in rows 9 through 13 above for columns A through E. Don't change anything in columns A through C. In column D, row 9, change the value of cell D9 to $=(A9-C15)^{0.25}$. Hence, you are simply changing the value of "0.5" to "0.25." Do the same for cells D10 through D13.

In column E, row 9, change the value of cell E9 to $=(1-B9)*A9^{0.25}+B9*0^{0.25}$. Again, you are simply changing "0.5" to "0.25." Do the same for cells E10 through E13. Then compare columns D and E.

PART 8

GOVERNMENT ACTIONS AND MANAGERIAL BEHAVIOR

CHAPTER 17

LEARNING OBJECTIVES

- Competition versus Monopoly
- Regulation of Monopoly
- The One Star Gas Company:
A Pseudo-Case Study
- Effects of Regulation on
Efficiency
- The Concentration of Economic
Power
- The Sherman Act
- The Clayton Act, the Robinson-
Patman Act, and the Federal
Trade Commission Act
- Interpretation of the Antitrust
Laws
- The Patent System
- Trade and Trade Policy
- Government Price Ceilings and
Price Floors
- The Welfare Impacts of Taxes
- Regulation of Environmental
Pollution
- Public Goods
- Summary
- Problems
- Excel Exercise: Externalities

GOVERNMENT AND BUSINESS

Bill Gates, chairman of Microsoft, the computer software giant, is a billionaire; yet he and other top executives of Microsoft had to be concerned with the U.S. Department of Justice's investigation of the firm's competitive practices. The government prevailed, and Microsoft was found guilty of antitrust violations. This set in motion appeals of the verdict. The government and Microsoft ultimately reached a compromise.

The government is a major player for business. Managers must be aware not only of antitrust laws but also of laws about fair trade, employment, safety, environmental issues, and securities, among others. U.S. companies, which often complain about excessive government interference with their activities, are now complaining about a lack of rules in the developing countries of the world. The game of business cannot be played without rules, and the government sets the rules.

In general, economic regulation has decreased in the United States in the last 30 years. Transportation and banking are two prime examples. Previously government agencies controlled which carriers could enter and exit the transportation industries, the prices they charged, and whether they could merge. In banking, the range of services banks could provide was once heavily regulated; banks could provide banking services but not insurance or brokerage. After deregulation, transportation became like the restaurant industry—any company could enter or exit and charge what it wished.

Yet noneconomic regulation (which can have major economic costs) has grown in recent years. For instance, safety regulations and hours of service (number

STRATEGY SESSION: Government Actions and the Financial Crisis of 2008

While monetary and fiscal policy is the subject of a course in macroeconomics, the nature of the financial crisis of 2008 causes us to make some comments regarding the role of government and its impact on managerial behavior.

In 2008, managers witnessed the virtual drying up of credit markets. This caused them to cut back on their investments, lay off workers, and reduce their estimates of future earnings. In 2009, managers were being squeezed on both the demand and supply side. We have emphasized that managers need to be prepared to manage during bust as well as boom times. So, what should efficient managers do?

Our answer is to follow the principles we discuss in this book. The economic environment follows a natural cycle where booms and busts must be anticipated. Managers need to be prepared for both. The optimization principles, the strategizing principles, and the information asymmetries don't change with the state of the economy; however, what ends up being a manager's best action likely does; but it still involves using the right tools to arrive at that decision.

What is the role of government as a result of the crisis? The first is monetary policy to get credit flowing so banks can lend money for working capital, so businesses have the cash to carry on their day-to-day operations (short-run) and can lend money for investment in the business (long-run). The second is fiscal policy, which is designed to invest in public infrastructure, create jobs, and generate a multiplier effect (with increases in consumption by those employed). The last is a major push to re-regulate the segments of the economy that are blamed for precipitating the financial crisis. We expected to see much tighter controls on financial markets, banks, investment companies, insurance companies, etc., to promote transparency (i.e., eliminate the asymmetries in information) and integrity.

The re-regulation is the result of a perceived market failure. When the financial system was significantly deregulated in the 1980s, policymakers (like Alan Greenspan) felt that financial markets would be self-regulating and that government oversight could be relaxed. The events of 2008 showed that markets can and do fail. Even Greenspan admitted he had too much trust in the ability of markets to regulate themselves. And it was basically the asymmetric information of Chapters 15 and 16 that led to the market failures.

We are now emerging from the crisis. The Dodd-Frank Wall Street Reform and Consumer Protection Act is overhauling the regulation of Wall Street. The Financial Crisis Inquiry Commission has put the blame of the crisis on "widespread failure in government regulation, corporate management, and heedless risk-taking by Wall Street," and the Securities and Exchange Commission and The Federal Reserve were chastised for not requiring the banks to hold more reserves to cover potential losses from the risky practices of negligent mortgage banking, and the risky security packages into which mortgages were placed. One claim cited was that the five largest investment banks held only \$1 in capital to cover losses for every \$40 of "assets" they held. So if the assets fell in value by 2.5%, that \$1 alone would be available to cover the losses. Because the "assets" in some cases were worthless, the \$1 didn't go a long way in covering a \$40 loss.

Source: Sewell Chan, "Financial Crisis Was Avoidable, Inquiry Finds," January 25, 2011, at www.nytimes.com/2011/01/26/business/economy/26inquiry.html.

of hours on duty) regulations continue in the transportation industry. Laws are passed each year that affect how business is conducted.

Governments collect taxes that affect disposable income and the final prices of goods. Governments also spend money that affects the demand for goods. Tax laws affect corporate investment policy. Governments subsidize certain goods, like agricultural products, to increase their production; and governments certify whether and when certain goods can appear in the market, as in the case of drugs (both legal and illegal). Governments provide infrastructure, like roads and water and sewer lines, which are an important part of many firms' production functions.

Thus government activity is pervasive. It accounts for approximately 30% of our gross domestic product. If you plan to move to the top of the executive ranks, be prepared to interact not only with your colleagues and rivals but also with government agencies.

The previous chapters have shown the power of markets to allocate resources and to get things done in the economy without any intervention. In general, the authors of this book believe that the market is a good solution. However, we also believe that markets can and do fail. Negative externalities such as congestion and pollution come to mind. In those circumstances a role for outside intervention may be called for, or the market must reinvent itself and come up with a new market mechanism that corrects the failure. Sometimes the government can nudge the market in this direction. Although externality markets initially called for intervention, market-based congestion tolls and tradable emission permits are becoming more prevalent. Thus new markets can evolve to solve previous market failures. Without a governmental nudge, however, the market failure might persist.

In this chapter we discuss how actions like public regulation, antitrust policy, trade policy, price intervention, taxes, and the patent system affect managerial life. In addition, we view the role of government in providing public goods and in correcting market failures. Managers must not only understand the nature of public policy in these areas; they must understand what public policy is designed to achieve. Too frequently business executives lack the breadth of view and knowledge required to effectively promote their firms' interests in the public arena (even though in 2011 there were 10,458 registered lobbyists, many of them corporate, in Washington—a number that has fallen by 17% since 2000).¹

We will leave it to your course in macroeconomics to discuss federal fiscal and monetary policy, both of which have major impacts on the economy and hence on your business.

COMPETITION VERSUS MONOPOLY

The Supreme Court has stated that competition is its fundamental national policy. Many economists agree that competition is preferable to monopoly (or other serious departures from perfect competition) because it is likely to result in a better

1. Lobbying Database at www.opensecrets.org/lobby/index.php. The number represents unique, registered lobbyists who have actively lobbied.

allocation of resources. As we saw in Chapter 8, a monopolist tends to restrict output, driving up the price. These economists argue that from the point of view of social welfare, it is better if a monopolist raises its output to the competitive level. (Also, in their view, monopolists are likely to be less efficient than competitive firms.) Although economists are not unanimous in this opinion, the majority seem to prefer competition over monopoly.

One way our society has dealt with these issues is to establish government commissions like the Federal Communications Commission and the Interstate Commerce Commission (now abolished but with some residual authority retained in the Surface Transportation Board) to regulate the behavior of monopolists. As we see in subsequent sections of this chapter, the government has tried to reduce the harmful effects of monopoly. In addition, Congress has enacted antitrust laws

STRATEGY SESSION: More Competition in Kenya

Safaricom is the largest telecom in Kenya. In 2010 it slashed its rates in half within days in response to its rivals slashing their rates “to the bone.” If Kenya’s proposed changes in their Information and Communication Act of 2008 had gone into effect, Safaricom would have had to wait 30 days before responding to the competition because they are defined as a dominant firm. Safaricom has an 80% market share, and dominance is defined as 25% (with a proposal that it be raised to 40%). Even so, in the several days it took for Safaricom to meet the rivals’ lower rates, it lost 2% of its market to competitor Zain. If the competitors had a 30-day head start on Safaricom, the damage would likely have been much higher. Zain’s parent company is Indian Bharti Airtel, and its marching orders from home are apparently to become the market leader within two years. Not being able to respond to a price cut by a rival if you are a dominant firm would seem to be a surefire way to bring up the market shares of the nondominant firms. But by how much? A report by Stanbic Investments states that “the Kenyan market is estimated to have a price elasticity of below 1.” Thus, the 50% decrease in rates would generate less than a 50% increase in business, and already the margins at Zain are described

as “very thin” and would only get thinner with further rate cuts.

MultiChoice Kenya is the dominant player in the pay-TV market in Kenya. But recently My TV Africa was launched, cable TV provider Zuku lowered its prices, and public broadcaster Kenya Broadcasting Corporation has partnered with Swedish Next Generation Broadcasting to offer Smart TV. All three of these companies provide competition to MultiChoice, and all three are charging rates below Kenyan Shilling (Ksh) 1,000, leading MultiChoice to cut its premium rate to Ksh 5,000 but also offer a package below Ksh 1,000. The country is switching to a digital format, which will lower cost and along with the lower rate increase demand for pay TV. The market is estimated to grow at 13% per year (the fastest growth in the world); however, the figure is misleading because it is starting from such a low base. But at that rate, the market could double in less than six years.

Source: Muna Wahome, “Kenya: Safaricom’s Hour of Reckoning,” *The Nation* (Nairobi), August 28, 2010, <http://allafrica.com/stories/201008300505.html>; Winfred Kagwe, “Kenya: Fight for Pay Television Market Down to the Wire,” *The Nation* (Nairobi), November 29, 2010, <http://allafrica.com/stories/201011300406.html>.

meant to promote competition and control monopoly. These laws too are discussed at length in this chapter. Any manager must be aware of the nature of these laws because violating them may trigger significant fines and jail sentences.

Until recently the United States went further in promoting competition than other major industrialized countries. But just because we promote competition does not mean our dedication to competition is complete. In the new millennium, the European Union seems to have taken the lead with respect to antitrust policy (in some cases pursuing cases that the United States has chosen not to prosecute). To some extent this represents a “catching up” because Europe historically has not pursued anticompetitive activities. Other countries, at our and the EU’s behest, such as South Korea, Japan, Brazil, China, and India, are now taking a more active role in antitrust enforcement. But national policies are too ambiguous and rich in contradictions to be characterized so simply. The truth is that we, as a nation, have adopted many measures to promote monopoly and limit competition. For example, this is the effect of the patent system, which is designed to promote invention and innovation. In later sections of this chapter we see why the patent system is beneficial even though it creates temporary monopolies.

REGULATION OF MONOPOLY

In some areas of the economy, such as the distribution of water, it is not economical for more than one firm to exist due to important economies of scale. In such industries, the single firm, a so-called natural monopolist, is in a position to charge a higher-than-competitive price for its product. Because such a price may inefficiently allocate society’s resources, as well as create monopolistic profit regarded by the public as excessive and unjustifiable, government regulatory commissions often are established to limit the prices a monopolist of this sort can charge.

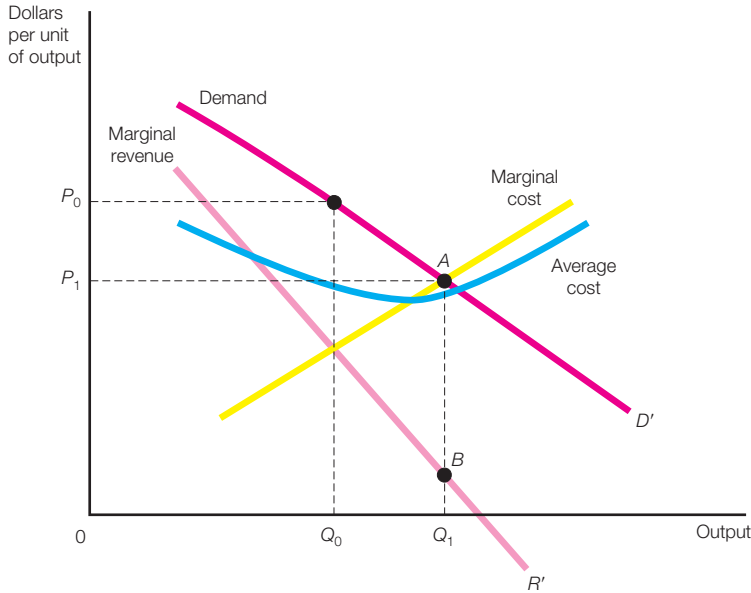
Consider the Acme Water Company, whose demand curve, marginal revenue curve, average cost curve, and marginal cost curve are shown in Figure 17.1. Without regulation, managers will charge a price of P_0 and produce Q_0 units of the product. By setting a maximum price of P_1 , the firm’s demand curve becomes P_1AD' , its marginal revenue curve becomes P_1ABR' , its optimum output becomes Q_1 , and it charges the maximum price of P_1 . By setting the maximum price, the commission aids consumers, who pay a lower price for more of the product. By the same token, the commission takes away some of the Acme Water Company’s monopoly power.

Regulatory commissions often establish the price (or the maximum price) at the level at which it equals average total cost, including a fair rate of return on the company’s investment. In California, this is defined as a cost-of-service or a rate-of-return model. The California Public Utilities Commission (PUC) examines the applicant’s costs in a test year and then determines the company’s revenue requirement during that test year. The PUC first determines the value of the assets

FIGURE 17.1

Regulation of Acme Water Company: Maximum Price

By setting a maximum price of P_1 , a regulatory commission can make Acme increase output to Q_1 .



the applicant has invested to provide its utility service. If an asset is found to be “unproductive” for public utility purposes, that asset is excluded from the applicant’s “rate base.” Then the applicant’s cost of capital necessary to invest in the assets in their rate base is calculated. The cost of capital times the rate base is a critical element in determining the applicant’s revenue need.

Debt and equity are the two ways for the applicant to raise capital. Debt is usually cheaper to obtain, but it increases the risk faced by the stockholders (equity providers) because the interest on the debt must be paid (before dividends can be paid) regardless of how well the applicant is doing financially. To adjust for that risk, shareholders expect a higher return on their investment. Most firms try to find a “middle ground” between these two methods of financing capital assets.

The PUC then determines the applicant’s cost of capital in three steps. In the first step, the PUC decides on a “reasonable” capital structure. This is a ratio of debt to equity that the PUC feels the applicant *should* use to finance its capital needs. Call this D/E . Adding E/E to this yields $(D + E)/E$ or C/E , where C is the applicant’s reasonable capital. The reciprocal of C/E is E/C , and that represents the reasonable share of equity in the applicant’s capital structure. Then $1 - (E/C) = D/E$ is the reasonable share of debt in the applicant’s financial

STRATEGY SESSION: The Social Cost of Monopoly

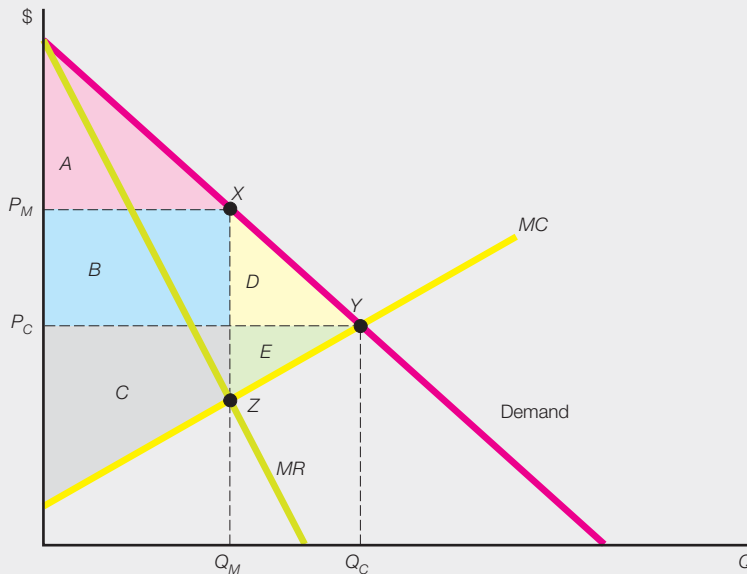
Consider the following figure. A monopolist would set price at P_M and output at Q_M . Consumer surplus under a monopoly would be A ; producer surplus under a monopoly would be $B + C$; hence the social welfare under a monopoly would be $A + B + C$.

If the market were perfectly competitive, price would equal marginal cost (P_C) and quantity would be $[Q_C]$. The consumer surplus under perfect competition would be $A + B + D$. The producer surplus would be $C + E$, and the social welfare $A + B + C + D + E$.

Therefore the social welfare under perfect competition is $D + E$ greater than under a monopoly.

This is often called the *social welfare triangle* or *dead-weight loss triangle*. The rationale for this welfare cost of monopoly is that the demanders along segment XY of the demand curve are willing to pay more than the marginal cost (ZY) of producing the goods (between the quantities $Q_C - Q_M$), yet the monopolist does not produce such socially beneficial goods (restricting output at Q_M).

Part of the rationale of antitrust policy and regulation is to ensure that society captures part or all of this $D + E$ triangle.

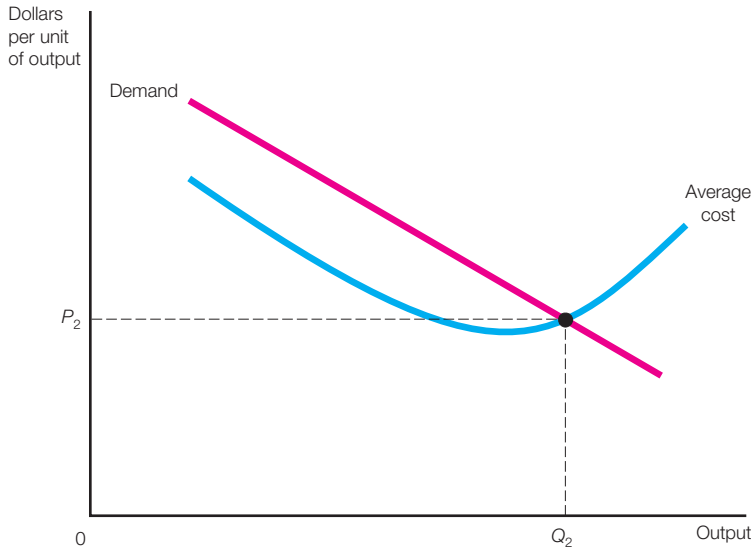


structure. In the second step, the PUC calculates the applicant's cost of debt based on the actual cost of the applicant's outstanding debt in the most recent period (i_D). In the third step, the PUC determines the appropriate return on the applicant's equity by looking at the returns on equity of other firms facing comparable risks (i_E). Applying the resulting costs of debt and equity capital to the applicant's reasonable capital structure shares yields the weighted average cost of capital i , that is, $i = (E/C)i_E + (D/C)i_D$.

FIGURE 17.2

Regulation of Acme Water Company: Fair Rate of Return

The regulated price is P_2 , where the demand curve intersects the average total cost curve, which includes what the commission regards as a fair profit per unit of output.



The PUC then multiplies the applicant's rate base by the weighted average cost of capital and adds the applicant's operating expenses and tax costs to determine the applicant's revenue requirement. This figure is used to determine rates that will cover the applicant's costs and provide a reasonable rate of return on its investments.²

Using the above method, in Figure 17.2 the price would be set by the commission at P_2 , where the demand curve intersects the average total cost curve. Although, the latter curve includes what the commission regards as a fair profit per unit of output, considerable controversy arises over what constitutes a fair rate of return, as well as what should be included in the company's investment on which the fair return is to be earned.

General rate cases play an important role in the regulatory process. Such cases are initiated by the firms, based on the firm's claims that earnings are too small and a higher price level is needed. As an example, the Golden State Water Company has submitted an application to the California PUC to have a rate of return on its rate base of 9.49% for the 2012 test year, 9.52% for the 2013 test year, and 9.55% for the 2014 attrition year. Its capital shares are 55.6% common equity and 44.4% long-term debt, and it argues that its return on equity should be 11.5%. It claims

2. See *In the Court of Appeal of the State of California, Fifth Appellate District, The Ponderosa Telephone Co. petitioner v. Public Utilities Commission*, F061287, CPUC Decision Nos. 10-06-029 & 10-10-036, Opinion, filed July 5, 2011, Background, "Ratemaking principles and procedures," pp. 3–4.

implementation of the 9.49% rate of return for 2012 will generate a \$6.2 million increase in revenue.³

In Wyoming, Montana-Dakota Utilities Co. has agreed to a 10.9% rate of return.⁴ In New Zealand, the Commerce Commission has proposed a weighted average cost of capital of around 8.8% that has left the country's largest network power provider, Vector, howling.⁵

Demand is generally assumed to be price-inelastic; consequently, higher prices are assumed to result in greater revenues. The industry generally receives less than it requests (and commission decisions lag behind the industry's revenue requests). However, the fact that the commission does not approve all requests does not imply that the company is constrained much by the commission—the company may have asked for more than it thought it would receive.

Commissions try to regulate the industry so that earnings equal a “reasonable return on the firm's existing plant.” Yet a host of questions arise concerning what constitutes a “reasonable return” and the “value of the existing plant.” The original or historical cost of the plant is the measure on which most commissions base their estimates of the value of the plant, but some permit firms to use replacement cost valuation instead. As mentioned above, boards look at a weighted cost of debt and risk-adjusted equity capital as a guideline to allowable returns. In many cases risk is determined by the beta from the capital asset pricing model.

THE ONE STAR GAS COMPANY: A PSEUDO-CASE STUDY

As an illustration of how regulatory commissions work, we look at a hypothetical case. Suppose in 2010 the One Star Gas Company, which provides gas to residential and commercial customers in a major Texas urban area, requests an increase in its price. The Railroad Commission of Texas is the state regulatory body with authority over gas company rates. To decide whether an increase should be granted, the commission begins by determining the appropriate rate base. The company's assets that are “used and useful” are identified and valued at their historical cost. After allowing for accumulated depreciation, the original cost of invested capital is calculated to be \$300 million.

To establish the rate of return that One Star Gas should earn on this invested capital, the cost of debt and equity capital and the percentage of each to total capitalization is estimated. Using percentages from a hypothetical example (found in the Natural Gas Rate Review Handbook, Gas Services Division, Railroad Commission of Texas, May 2010, Table iii–5, p. 30), the cost of capital is 10.50% for common equity and 7.25% for long-term debt. These two costs are weighted by the share of the total capitalization of One Star Gas (45% for common equity and 55% for long-term debt). The result is $7.25 \times 0.55 + 10.50 \times 0.45 = 8.7125\%$, that is, the weighted average cost of capital for One Star Gas is 8.7125%. It should be noted that many firms have preferred equity (preferred stock in addition to com-

3. See “Before The Public Utilities Commission of the State of California, Application of Golden State Water Company (U 133 W) for Authority to Establish its Authorized Cost of Capital and Rate of Return for Utility Operations for 2012–2014,” Application 11-05, May 2, 2011, pp. 6–7.

4. Carl Shaw, “Utility Rates in Wyoming Predicted to Rise Dramatically,” November 4, 2010, at www.utilitiesanalyses.com/resources/Utility_Rates_in_Wyoming_Predicted_to_Rise_dramatically.php.

5. “Electricity Sector: Networks Cry Foul Over Regulated Rates of Return,” October 27, 2010, at www.nzenergy-environment.co.nz/home/free-articles.

PROBLEM SOLVED: The Trenton Gas Company

To illustrate the workings of public utility regulation, consider the Trenton Gas Company, which has assets of \$300 million. The state Public Utility Commission, after considering the extent of the risks assumed by the firm and conditions in the financial markets, decides that a fair rate of return for the firm would be 10%. Trenton Gas is allowed to earn a profit of $0.10(\$300 \text{ million}) = \30 million per year. This profit is not economic profit; it is an accounting profit. As stressed already, commissions try to allow firms to earn only a normal or fair rate of return, which tends to rule out economic profit.

What price and output will managers choose? To answer these questions we note that the demand curve for gas provided by the firm is

$$P = 30 - 0.1Q \quad (17.1)$$

where P is the price per unit of gas (in dollars) and Q is the number of units of gas demanded (in millions). The firm's total cost equals

$$TC = 10 + 5Q + 0.9Q^2 \quad (17.2)$$

where TC is total cost (in millions of dollars). Note that this concept of total cost does not include the opportunity cost of the capital invested in the firm by its owners. Therefore, the firm's accounting profit equals

$$\begin{aligned} \pi &= (30 - 0.1Q)Q - (10 + 5Q + 0.9Q^2) \\ &= -Q^2 + 25Q - 10 \end{aligned} \quad (17.3)$$

where π is the firm's profit (in millions of dollars).

Because the commission has decided that the firm's accounting profit should equal \$30 million, we set π equal to 30, which implies that

$$\begin{aligned} 30 &= -Q^2 + 25Q - 10 \\ Q^2 - 25Q + 40 &= 0 \end{aligned} \quad (17.4)$$

which is an equation of the form $aQ^2 + bQ + c = 0$. We can use the following equation to solve for the roots of this equation

$$\begin{aligned} Q &= \frac{-b \pm (b^2 - 4ac)^{0.5}}{2a} \\ &= \frac{25 \pm [(-25)^2 - 4(1)(40)]^{0.5}}{2(1)} \\ &= \frac{25 \pm (465)^{0.5}}{2} \\ &= \frac{25 \pm 21.56}{2} \\ &= 1.72 \text{ or } 23.28 \end{aligned}$$

Because commissions generally want public utilities to serve as many customers as possible, the larger figure, $Q = 23.28$, is the relevant one. Therefore, the price is set as follows

$$\begin{aligned} P &= 30 - 0.1(23.28) \\ &= 27.67 \end{aligned}$$

To sum up, the Trenton Gas Company's price will be \$27.67, and it will deliver 23.28 million units of gas.

mon stock), and short-term debt in addition to long-term debt. In this case, the cost of capital of all four classes of capital would be determined and then weighted by their shares of total capitalization to yield a weighted average cost of capital. In the example at hand, the commission would conclude that One Star Gas should earn an 8.7125% return on its invested capital of \$300 million.

STRATEGY SESSION: A Dispute over a Requested Gas Rate Increase

The Boston Gas Company's request for a \$17 million rate increase was examined at hearings conducted by the Massachusetts Department of Public Utilities. One consumer advocacy group, which opposed the rate increase, argued that the gas company should be allowed a 10.5% rate of return, whereas the company asked for a 12.46% rate of return. The company also argued that because of the regulatory lag, it was receiving considerably less than the 12% return it was allowed by the commission. Because the previous rate increases had not become effective until almost a year after they were requested, the company earned about 9%, not 12%.

At the hearing an economist testifying for the consumer advocacy group argued that the company's cost of equity capital was about 12%, whereas an economist hired by the company testified that it was about 16%. This was an important issue in the case. Both economists used the following equation:

$$k_t = \frac{D_1}{W} + g$$

Here k_t is the required cost of capital at time t , D_1 is the dividend paid next year, W is the firm's capitalized value, and g is the annual rate of growth in the corporate dividend, to estimate the company's cost of equity capital. The economist hired by the consumer advocacy group assumed that the annual rate of dividend growth, g , would equal 0.01, whereas the company's economist assumed that it would equal 0.05.

If asked to advise this firm, what suggestions would you make concerning ways to reduce the adverse effects of the regulatory lag on the company's earnings? What sorts of analyses would you carry out to determine which of the two estimates of the cost of equity capital is closer to the truth?^a

^aFor further discussion, see Barrett and Cormack, *Management Strategy in the Oil and Gas Industries*.

To earn this return, One Star should make an annual profit of $0.087125 \times \$300$ million, or \$26.1375 million. Suppose One Star's actual profit is \$20 million. Because One Star's profit is \$20 million, the commission would likely decide to allow One Star to raise its price to bring its profit up to \$26.1375 million.

EFFECTS OF REGULATION ON EFFICIENCY

Regulators try to prevent a monopoly from earning excessive profits. As we saw in previous sections, managers are permitted only a fair rate of return on their investment. A difficulty with this arrangement is that they are guaranteed this rate of return, regardless of how poorly they manage. If the regulators decide the Acme Water Company should get a 9% rate of return on its investment, it will get this rate of return regardless of whether it is managed well or poorly. Why is this a problem? Because unlike a competitive firm, there is no incentive for managers to increase firm efficiency.

Managers need to anticipate a regulatory process that is characterized by long delays. If they are smart, managers will use these delays to increase firm efficiency. In many regulated industries, a proposed rate change may be under review for months before a decision is made by the commission. If such a price change is hotly contested, it may take years for the required hearings and subsequent court appeals. Such a delay between a proposed price change and its ultimate disposition is called a **regulatory lag**. Long regulatory lags are often criticized by people who would like the regulatory process to adapt more quickly to changing conditions and yield more timely decisions. But an advantage of regulatory lags is that they penalize inefficiency and reward efficiency. To combat the problem of regulatory lag, some statutes require regulatory agencies to render decisions within a specific time period.

Regulatory lag A delay between a proposed price change and its ultimate disposition.

To illustrate, consider managers at a regulated company whose price is established so they can earn a 9% rate of return (which the commission regards as a fair rate of return). The managers develop and introduce an improved manufacturing process that cuts costs, thus increasing their actual rate of return to 11%. If it takes 15 months for the commission to review the prices it approved before and modify them to take account of the new (lower) cost levels, managers can earn a higher rate of return (11% rather than 9%) during those 15 months.

Although the regulatory lag restores some incentives for efficiency (and some penalties for inefficiency), it does not result in as strong a set of efficiency incentives as competitive markets. A fundamental problem with regulation is that if a regulatory commission prevents a firm from earning higher-than-average profits, there may be little incentive for managers to increase efficiency and introduce innovations.

THE CONCENTRATION OF ECONOMIC POWER

Government regulatory commissions are not the only device used by society to deal with the problem of monopoly; another device is antitrust laws. These laws reflect a feeling that excessive power lies in the hands of relatively few firms. According to the *2011 Statistical Abstract of the United States*, manufacturing firms with more than \$250,000 in assets had a total book asset value of over \$7.0 trillion in 2009. According to the latest *Census of Manufacturing* (2007), the 100 largest manufacturing firms employed 15.2% of the manufacturing workforce, paid 20.8% of the manufacturing payroll, used 14.2% of manufacturing production workers, created 34.6% of the value added in manufacturing, purchased 37.4% of the raw materials used in manufacturing, made 33.8% of all the capital expenditures by manufacturing firms, and had 36.1% of the total value of shipment shipped. These percentages have increased considerably since the end of World War II. Although bigness is not necessarily the same as monopoly power, there is a widespread feeling that economic power is concentrated in relatively few hands.

Antitrust laws are aimed at promoting competition and limiting monopoly. As mentioned earlier, many economists believe that competition is preferable to monopoly because competition tends to more effectively allocate resources. To measure how close a particular industry is to being perfectly competitive (or monopolized), economists have devised the *market four-firm concentration ratio*, which shows the percentage of total sales or production accounted for by the industry's four largest firms. The higher the percentage, the more concentrated the industry.

Table 17.1 shows market four-firm concentration ratios for selected non-manufacturing industries—that is, the market share of the largest four firms. For each industry sector, the highest ratio, the lowest ratio, and the ratio of the subsector with the highest revenue of the sector are reported. These ratios vary widely from industry to industry. In the discount department store industry, the concentration ratio is very high: 96.9%. In the chiropractors industry, it is very low: 0.6%. The concentration ratio is only a rough measure of an industry's market structure, which must be supplemented with data on the extent and type of product differentiation in the industry, as well as on barriers to entry. Even with these supplements, it is still a crude measure because, for one thing, it takes no account of competition from foreign suppliers. Nonetheless, the concentration ratio has proven to be a valuable tool.

Another measure of concentration shown in Table 17.1 is the **Herfindahl-Hirschman index (HHI)**, which equals the sum of the squared market shares of all the firms in the market for manufacturing industries. This index is used by the U.S. Department of Justice, Antitrust Division, and the Federal Trade Commission as a guide to determine whether they should investigate proposed mergers.⁶ For example, if two firms exist in a market and each has a 50% share of the market, this index equals $50^2 + 50^2 = 5,000$. The HHI can range from 10,000 (a monopolist would have an HHI of 100^2) to 0 (an infinite number of atomistic competitors each with a market share approaching 0). According to the government's merger guideline, if the HHI (after the merger) will be less than 1,000, the merger is unlikely to be challenged (but this is only a guideline, so nothing is guaranteed). If the postmerger HHI will be between 1,000 and 1,800 and the index changes by less than 100 points as a result of the merger, the merger is unlikely to be challenged. Finally, if the postmerger HHI will be greater than 1,800 and the index changes by less than 50 points as a result of the merger, the government is unlikely to challenge the merger. Note that this last requirement may be hard to satisfy. For instance, if a firm with a 49% share wanted to merge with a firm with a 1% share, the HHI would increase by 99 points. A firm with a 25% share merging with a firm with a 1% share would raise the HHI by 51 points. Firms with low four-firm concentration ratios tend to have low HHIs (the HHI is not calculated for all firms here but rather for the largest 50—thus the HHI is understated). For instance, the HHI for the machine shops industry is just 2.6, whereas the HHI for the fiber can, tube, and drum industry is 2,874.1. Note that certain industries, such as the

Herfindahl-Hirschman index (HHI) An index that equals the sum of the squared market shares of all the firms in the market of manufacturing industries.

6. See U.S. Department of Justice, Antitrust Division, *1997 Merger Guidelines*, at www.usdoj.gov/atr/hmerger/11251.htm.

TABLE 17.1

Concentration Ratios and Herfindahl-Hirschman Indexes (HHI) by Economic Sectors for Largest (by Revenues or Value of Shipments) Subsector for Each Three-Digit NAICS Sector (Bold) and by Largest and Smallest Four Firm Concentration Ratio or HHI for Each Sector: 2007

	Market Share of Four Largest Firms	Herfindahl-Hirschman Index for 50 Largest Firms
NAICS Sector		
22 Utilities		
221122 Electric Power Distribution	17.7	
221330 Steam and Air Conditioning Supply	61.4	
311 Food Manufacturing		
311312 Cane Sugar Refining	95.2	Not Given
311612 Animal (Except Poultry) Slaughtering	59.4	1,046.5
311811 Retail Bakeries	3.7	7.3
312 Beverage and Tobacco Product Manufacturing		
312140 Distilleries	42.3	584.0
312221 Cigarette Manufacturing	97.8	Not Given
313 Textile Mills		
313113 Thread Mills	65.6	1,828.7
313230 Nonwoven Fabric Mills	45.3	778.2
313311 Broadwoven Fabric Finishing Mills	21.2	274.8
314 Textile Mill Products		
314110 Carpet and Rug Mills	63.6	1,649.7
314912 Canvas and Related Product Mills	12.6	90.6
314992 Tire Core and Tire Fabric Mill	72.0	Not Given
315 Apparel Manufacturing		
315192 Underwear and Nightwear Knitting Mills	98.1	D
315212 Women's, Girl's, and Infant's Cut and Sew Apparel Contractors	6.2	19.2
316 Leather and Allied Product Manufacturing		
316110 Leather and Hide Tanning and Finishing	38.5	560.3
316214 Women's Footwear (Except Athletic) Manufacturing	77.8	2,543.8
316991 Luggage Manufacturing	32.2	399.8
321 Wood Product Manufacturing		
321113 Sawmills	14.6	98.3
321212 Softwood Veneer and Plywood Manufacturing	55.7	1,233.3
321999 All Other Miscellaneous Wood Product Manufacturing	10.9	65.1
322 Paper Manufacturing		
322121 Paper (Except Newsprint) Mills	49.8	758.9
322214 Fiber Can, Tube, Drum, and Similar Products Manufacturing	77.7	2,874.1

322299 All Other Converted Paper Product Manufacturing	20.5	187.7
323 Printing and Related Support Activities		
323111 Commercial Gravure Printing	65.2	1,359.5
323113 Commercial Screen Printing	12.2	59.4
323114 Quick Printing	3.8	8.4
324 Petroleum and Coal Products Manufacturing		
324110 Petroleum Refineries	47.5	806.5
324121 Asphalt Paving Mixture and Block Manufacturing	21.8	188.4
324122 Asphalt Shingles and Coating Materials Manufacturing	67.0	1,257.6
325 Chemical Manufacturing		
325192 Cyclic Crude and Intermediate Manufacturing	89.7	2,328.9
325412 Pharmaceutical Preparation Manufacturing	34.5	456.8
325008 Other Miscellaneous Chemical Product Manufacturing	18.9	154.8
326 Plastics and Rubber Products Manufacturing		
326199 All Other Plastics Product Manufacturing	6.3	26.3
326211 Tire Manufacturing (Except Retreading)	77.6	1,734.7
327 Nonmetallic Mineral Product Manufacturing		
327111 Vitreous China Plumbing Fixture Manufacturing	87.2	D
327320 Ready-Mix Concrete Manufacturing	22.6	312.9
327991 Cut Stone and Stone Product Manufacturing	6.9	22.6
331 Primary Metal Manufacturing		
331111 Iron and Steel Mills	52.1	785.6
331423 Secondary Smelting, Refining and Alloying of Copper	63.5	1,217.4
331524 Aluminum Foundries (Except Die-Casting)	27.5	288.8
332 Fabricated Metal Product Manufacturing		
332710 Machine Shops	1.7	2.6
332993 Ammunition (Except Small Arms Manufacturing)	8.2	2,015.9
333 Machinery Manufacturing		
333120 Construction Machinery Manufacturing	53.6	1,143.3
333511 Industrial Mold Manufacturing	4.6	21.0
333913 Measuring and Dispensing Pump Manufacturing	72.7	1,729.3
334 Computer and Electronic Product Manufacturing		
334112 Computer Storage Device Manufacturing	86.9	D
334413 Semiconductor and Related Device Manufacturing	55.7	1,283.6
334419 Other Electronic Component Manufacturing	11.7	75.0
335 Electrical Equipment, Appliance, and Component Manufacturing		
335222 Household Refrigerator and Home Freezer Manufacturing	91.6	Not Given
335929 Other Communication and Energy Wire Manufacturing	36.6	479.8
335999 Miscellaneous Electrical Equipment Manufacturing	19.6	165.9
336 Transportation Equipment		
336112 Light Truck and Utility Vehicle Manufacturing	84.3	2,679.5
336211 Motor Vehicle Body Manufacturing	23.6	243.0

THE CONCENTRATION OF ECONOMIC POWER

335414 Guided Missile and Space Vehicle Manufacturing	94.6	Not Given
337 Furniture and Related Product Manufacturing		
337121 Upholstered Household Furniture Manufacturing	34.0	375.5
337125 Household Furniture (Except Wood and Metal)		
Manufacturing	67.0	2,215.4
335127 Institutional Furniture Manufacturing	13.1	87.8
339 Miscellaneous Manufacturing		
339950 Sign Manufacturing	6.7	27.9
339995 Burial Caskets Manufacturing	73.5	1,979.1
3399999 All Other Miscellaneous Manufacturing	26.2	274.5
423 Merchant Wholesalers Durable Goods		
4234301 Computer and Computer Peripheral Equipment	33.0	
4238201 Farm Machinery and Equipment	4.1	
42399042 Prerecorded Video Tapes	78.8	
424 Merchant Wholesalers Non-Durable Goods		
4242101 General-Line Drugs	74.6	
4244905 Food and Beverage Basic Material	87.0	
4249901 Art Goods	5.7	
425 Wholesale Electronic Markets, Agents and Brokers		
42512047 Petroleum and Petroleum Products	76.1	
42512031 Motor Vehicle and Motor Vehicle Parts	70.0	
42512036 Electrical and Electronic Goods	4.5	
441 Motor Vehicle and Parts Dealers		
441110 New Car Dealers	6.0	
441221 Motorcycle, ATV, and Personal Watercraft Dealers	3.7	
441310 Automotive Parts and Accessories Stores	34.8	
442 Furniture and Home Furnishings Stores		
442210 Floor Covering Stores	2.9	
442299 All Other Home Furnishings Stores	48.8	
443 Electronics and Appliance Stores		
443111 Household Appliance Stores	21.3	
443112 Radio, Television and Other Electronics Stores	70.0	
443120 Computer and Software Stores	73.1	
444 Building Materials and Garden Equipment and Supplies Dealers		
444110 Home Centers	92.7	
444210 Outdoor Power Equipment Stores	4.7	
445 Food and Beverage Stores		
445110 Supermarkets and Other Grocery (Except Convenience) Stores	32.0	
445220 Fish and Seafood Markets	4.1	
445292 Confectionary and Nut Stores	36.9	
446 Health and Personal Care Stores		

446110 Pharmacies and Drug Stores	63.0
446199 All Other Health and Personal Care Stores	12.1
447 Gasoline Stations	
447110 Gasoline Stations With Convenience Stores	11.4
447190 Other Gasoline Stations	30.5
448 Clothing and Clothing Accessory Stores	
448120 Women's Clothing Stores	23.7
448140 Family Clothing Stores	45.4
4482103 Children's and Juveniles' Shoe Stores	78.2
451 Sporting Goods, Hobby, and Musical Instrument Stores	
4511101 General-Line Sporting Goods Stores	52.6
4511102 Specialty-Line Sporting Goods Stores	19.0
4512113 College Book Stores	74.8
452 General Merchandise Stores	
452112 Discount Department Stores	96.9
452910 Warehouse Clubs and Supercenters	93.9
4529904 Miscellaneous General Merchandise Stores	40.8
453 Miscellaneous Stores Retailers	
453110 Florists	2.1
453210 Office Supplies and Stationery Stores	80.3
454 Nonstore Retailers	
454111 Electronic Shopping	D
454112 Electronic Auctions	D
454113 Mailorder Houses	33.9
454311 Heating Oil Dealers	11.9
454312 Liquefied Petroleum Gas (Bottled Gas) Dealers	36.9
48 and 49 Transportation and Warehousing	
481111 Scheduled Passenger Air Transportation	52.3
484220 Specialized Freight (Except Used Goods) Trucking Local	2.6
488111 Air Traffic Control	94.4
51 Information	
511191 Greeting Card Publishers	90.9
512240 Sound Recording Studios	9.8
517110 Wired Telecommunications Carriers	56.8
52 Finance and Insurance	
522110 Commercial Banking	31.8
522294 Secondary Market Financing	97.9
522310 Mortgage and Nonmortgage Loan Brokers	5.2
53 Real Estate and Rental and Leasing	
531120 Lessors of Non-Residential Buildings (Except Miniwarehouses)	14.3
531312 Nonresidential Property Manager	6.7

THE CONCENTRATION OF ECONOMIC POWER

532111 Passenger Car Rental	82.0
54 Professional, Scientific, and Technical Services	
541110 Offices of Lawyers	2.6
541430 Graphic Design Services	1.8
541513 Computer Facilities Management Services	63.5
56 Administrative and Support and Waste Management and Remediation Services	
561110 Office Administrative Services	2.3
561320 Temporary Help Services	15.6
562213 Solid Waste Combustors and Incinerators	92.4
61 Educational Services	
611512 Flight Training	52.0
611610 Fine Arts Schools	3.2
611710 Educational Support Services	21.2
62 Health Care and Social Assistance	
621310 Offices of Chiropractors	0.6
621491 HMO Medical Centers	88.3
622110 General Medical and Surgical Hospitals	7.8
71 Arts, Entertainment, and Recreation	
713110 Amusement and Theme Parks	70.7
713930 Marinas	3.0
713940 Fitness and Recreational Sports Centers	15.5
72 Accommodation and Food Services	
722110 Full Service Restaurants	9.3
722310 Food Service Contractors	68.2
722320 Caterers	2.3
81 Other Services (Except Public Administration)	
812320 Drycleaning and Laundry Services (Except Coin Operated)	1.6
812332 Industrial Laundries	71.6
813211 Grant Making Foundations	8.4

The NAICS sector numbers refer to industries classified by the North American Classification System, a system used by Canada, Mexico, and the United States.

HHIs are only given for the Manufacturing Sectors, 31–33.

D = Not Disclosed because of Census Bureau disclosure rules.

Not Given = no HHI was given by the Census Bureau.

Source: Concentration Ratios: 2007. U.S. Census Bureau, U.S. Department of Commerce, Economics and Statistics Administration at <http://factfinder.census.gov>, various tables.

STRATEGY SESSION: The Art of the Deal Is Back

The year 2011 started off as the biggest year for takeovers since 2007. Acquiring firms paid a median of 9.2 times EBITDA (earnings before interest, taxes, depreciation, and amortization). This is still below the 11.4 times EBITDA paid in 2008, but it's the volume of takeovers that is making the news. Firms are buying rather than making to outpace their rivals. Duke Energy and Deutsche Boerse (merger with NYSE Euronext) are examples of acquirers. The activity is not sector specific, which indicates that the growth in the economy is robust across industries.

Since the Great Recession started, a number of firms have been making record profits but sitting on the cash they earned. The equity markets are now pushing managers not to hoard cash but to make bold strategic moves. Managers will be evaluated not based on mere accumulation of cash but based on how they invest this cash.

Estimates of cash holdings of up to \$3.3 trillion exist globally. The experts feel there is pent-up demand out there, and so far the Arab Spring and the Japanese tsunami of 2011 have not derailed the enthusiasm for the deal. Market activists such as Carl Icahn and hedge fund and equity capital firms are playing a role in activating managers. Firm managers can hear their stakeholders clamoring for higher returns, and they can hear market activists telling the shareholders that activists can deliver if current management cannot.

Source: Zachary R. Mider, "New Deal Rush Pushes Takeovers to Most Expensive Since Lehman," *Bloomberg Businessweek*, March 30, 2011, at www.businessweek.com/news/2011-03-30/deal-rush-pushes-takeovers-to-most-expensive-since-lehman.html.

cigarette and underwear and nightwear knitting mills (with four-firm concentration ratios of 97.8 and 98.1, respectively) industries, have no published HHIs. This is because the government has rules about disclosing information about firms. If both the four-firm concentration ratio and the HHI were given, clever people could discern which firms had what share (and hence what sales levels). With so few firms in each of these industries, we can presume that their 50-firm HHIs are quite high (if not 10,000).

THE SHERMAN ACT

The first federal antitrust law, the Sherman Act, was passed by Congress in 1890. Although the common law had long outlawed monopolistic practices, it seemed to many Americans in the latter part of the 19th century that legislation was needed to discourage monopoly and to preserve and encourage competition. The formation of trusts (monopolistic combines that colluded to raise prices and restrict output) brought the matter to a head. The essence of the Sherman Act lies in the following two sections:

Sec. 1. Every contract, combination in the form of trust or otherwise, or conspiracy, in restraint of trade or commerce among the several states or with foreign nations, is hereby declared to be illegal. Every person who shall make any such contract or engage in any such combination or conspiracy, shall be deemed guilty of a misdemeanor. . . .

Sec. 2. Every person who shall monopolize, or attempt to monopolize or combine or conspire with any other person or persons, to monopolize any part of the trade or commerce among the several States, or with foreign nations shall be deemed guilty of a misdemeanor.

In 1974 the Sherman Act was amended, making violations felonies rather than misdemeanors. Corporations can now be fined up to \$100 million, and individuals can be fined up to \$1 million and receive prison terms of up to 10 years. In addition to criminal fines and jail sentences, firms and individuals can be sued for triple damages in civil suits brought by those hurt by an antitrust violation.

It is important to recognize that if executives of two or more firms in a particular industry talk about prices and agree to fix them, they are in violation of Section 1 of the Sherman Act. To illustrate this point, consider Robert Crandall, former chief executive officer of American Airlines. He called Howard Putnam, then chief executive officer of Braniff Airways, on February 21, 1982, and proposed that they raise prices. The telephone call, which (unknown to Crandall) was taped, went as follows:

Putnam: Do you have a suggestion for me?

Crandall: Yes, I have a suggestion for you. Raise your goddamn fares 20 percent. I'll raise mine the next morning.

Putnam: Robert, we . . .

Crandall: You'll make more money and I will, too.

Putnam: We can't talk about pricing!

Crandall: Oh [expletive deleted], Howard. We can talk about any goddamn thing we want to talk about.⁷

After finding out about the call, the Justice Department filed a suit accusing Robert Crandall of breaking the antitrust laws by proposing to fix prices. But because there had been no agreement to fix prices, Section 1 had not been violated. Nonetheless the court decided that a proposal of this sort could be an attempt to monopolize part of the airline industry, which is forbidden by Section 2 of the Sherman Act. American Airlines said that it would not do such a thing again.

7. *New York Times*, February 24, 1983; www.nytimes.com/1983/02/24/business/blunt-talk-on-the-phone.html.

THE CLAYTON ACT, THE ROBINSON-PATMAN ACT, AND THE FEDERAL TRADE COMMISSION ACT

During its first 20 years, the Sherman Act was not regarded by its supporters as being very effective. The ineffectiveness of the Sherman Act led Congress in 1914 to pass two other laws: the Clayton Act and the Federal Trade Commission Act. The Clayton Act attempted to be more specific than the Sherman Act in identifying certain practices that were illegal because they would “substantially lessen competition or tend to create a monopoly.”

The Clayton Act outlawed unjustified price discrimination, which (as you recall from Chapter 9) is a practice whereby one buyer is charged more than another buyer for the same product. However, discrimination resulting from differences in the quality or quantity of the product sold or resulting from differences in cost or competitive pressures was allowed. In 1936 the Robinson-Patman Act amended the Clayton Act. It prohibited charging different prices to different purchasers of “goods of like grade and quality” where the effect “may be substantially to lessen competition or tend to create a monopoly in any line of commerce, or to injure, destroy, or prevent competition with any person who either grants or knowingly receives the benefit of such discrimination, or with customers of either of them.” The Robinson-Patman Act was aimed at preventing price discrimination in favor of chain stores that buy goods in large quantities. Small independent retailers felt threatened by the chain stores and pushed hard for this law.

The Clayton Act also outlawed the use of tying contracts that reduce competition. As Chapter 10 indicated, tying contracts make buyers purchase other items to get the product they want. For a long time IBM rented, but would not sell, its machines and insisted that customers must buy IBM ancillary equipment and use IBM maintenance services. The Supreme Court required IBM to end its tying contracts. However, not all tying contracts have been prohibited. If a firm needs to maintain control over complementary goods and services to make sure its product works properly, this can be an adequate justification for a tying contract. Also, if the tying arrangements are voluntary and informal, there is no violation of the law. Thus if a customer bought IBM ancillary equipment because that firm felt that it worked best on IBM primary equipment, this was no violation of the law so long as this customer did not have to buy IBM ancillary equipment. McDonald’s requires that its franchisees buy certain products from McDonald’s or from McDonald’s-approved vendors. The reason is that the value of one McDonald’s franchise depends on the quality of service provided by all franchises. Thus McDonald’s assures the reputation of all McDonald’s franchises by tying agreements.

Further, the Clayton Act outlawed mergers that substantially lessen competition; but because it did not prohibit one firm purchasing a competitor’s plant and equipment, it really could not stop mergers. In 1950 this loophole was closed by

the Celler-Kefauver Antimerger Act. However, this does not mean that mergers have become less prevalent. On the contrary, an epidemic of mergers in the 1980s continues to this day.

The Federal Trade Commission Act was aimed at preventing undesirable and unfair competitive practices. It established the Federal Trade Commission to investigate unfair and predatory practices and issue cease-and-desist orders. The act stated that “unfair methods of competition in commerce are hereby declared unlawful.” The commission—composed of five commissioners, each appointed by the president for a term of seven years—was given the formidable task of defining exactly what was “unfair.” Eventually the courts took away much of the commission’s power; but in 1938 the commission acquired the function of outlawing untrue and deceptive advertising. Also, the commission has authority to investigate various aspects of the structure of the U.S. economy.

INTERPRETATION OF THE ANTITRUST LAWS

The real impact of the antitrust laws depends on how the courts interpret them, and the judicial interpretation of these laws has varied substantially from one period to another. Typically charges are brought against a firm or group of firms by the Antitrust Division of the Department of Justice; a trial is held; and a decision is reached by the judge. In major cases, appeals are made that eventually could reach the Supreme Court.

In 1911, as a consequence of the first major set of antitrust cases, the Standard Oil Company and the American Tobacco Company were forced to give up a large share of their holdings of other firms. The Supreme Court, in deciding these cases, put forth and used the famous **rule of reason**—that only unreasonable combinations in restraint of trade, not all trusts, required conviction under the Sherman Act. In 1920 the rule of reason was employed by the Supreme Court in its finding that U.S. Steel had not violated the antitrust laws even though it had tried to monopolize the industry because the court said the company had not succeeded. U.S. Steel’s large size and its potential monopoly power were ruled beside the point because “the law does not make mere size an offense. It . . . requires overt acts.”

In the 1920s and 1930s the courts, including the conservative Supreme Court, interpreted the antitrust laws in such a way that they had little impact. Although Eastman Kodak and International Harvester controlled substantial shares of their markets, the Court, using the rule of reason, found them innocent on the grounds that they had not built up their near-monopoly positions through overt coercion or predatory practices.

During the late 1930s this situation changed dramatically with the prosecution of the Aluminum Company of America (Alcoa). This case, decided in 1945 (but begun in 1937), reversed the decisions in the U.S. Steel and International Harvester cases. Alcoa achieved its 90% of the market by means that would have

Rule of reason Rule stating that only unreasonable combinations in restraint of trade, not all trusts, required conviction under the Sherman Act.

STRATEGY SESSION: Antitrust Violations

All defendants should have their day in court to defend themselves, and everyone is innocent until proven guilty. The accusations of antitrust violations keep piling up, either brought on by an aggrieved party against the supposed aggriever or by the government antitrust authorities against the alleged violator.

Dean Foods Co. is the largest U.S. milk products producer. The Dairy Farmers of America is the largest cooperative organization representing dairy farmers in the United States. Both were sued by a group of dairy farmers in the northeastern United States. In addition, HP Hood and Dairy Marketing Services, an affiliate of the Dairy Farmers of America, were also named as defendants. Dean and Hood bottle about 90% of the milk in the Northeast.

The farmers accused the defendants of operating a monopoly (really a monopsony). The farmers claimed that they receive lower prices for their milk because Dean buys exclusively from the Dairy Farmers of America Cooperative and that relationship forces the independent dairy farmers to join the cooperative to survive. The case alleges that Dean

and the cooperative fix prices through various contracts, agreements, and understandings in violation of federal and state antitrust law. They have merged, acquired, and closed down bottling plants to gain that 90% market share.

The defendants claim that their actions have cut costs and increased efficiency, and that their actions are designed to increase the price paid to dairy farmers as well as their net returns. The suing dairy farmers claimed that the only way they can get access to the bottling plants is to pay dues to the Dairy Farmers of America or to Dairy Marketing Services and that they must do so to receive minimum monthly payments on grade A milk sales set by the USDA (U.S. Department of Agriculture). This compelled them to deal with the monopoly.

Source: "Dean Foods, Dairy Co-op Sued Over Alleged Monopoly," Bloomberg.co.jp, October 9, 2009, at www.bloomberg.co.jp/apps/news?pid=90970900&sid=aWhpmLkcxWWE.

been regarded as "reasonable" in earlier cases: Keeping its price low enough to discourage entry, adding capacity to take care of increases in the market, and so forth. Nonetheless the court decided that Alcoa, because it controlled practically all the industry's output, violated the antitrust laws.

Frustrating as it sometimes may be to managers, the antitrust laws are rather vague and ambiguous; consequently it is not easy to tell whether certain actions are permissible. Take the case of two breweries, Pabst and Blatz, that wanted to merge in 1958. The government objected to this merger even though the two firms together accounted for less than 5% of the nation's beer sales. What troubled the government was that they accounted for about 24% of beer sales in Wisconsin. The district court judge, agreeing with Pabst and Blatz that Wisconsin should be viewed as only part of the relevant market, dismissed the complaint; but the Supreme Court ruled against the firms. This case shows how difficult it can be to establish even the boundaries of the relevant market.

STRATEGY SESSION: China: New to Antitrust

China's National Development and Reform Commission (their top economic-planning agency) released their formulation of the country's rules on price monopolies in January 2011. This move is portrayed as a step in the enforcement of China's antitrust laws, which were put into place only in 2008. According to the new rules, price-fixing agreements are banned, and excluding or limiting competition through abuse of a dominant market share is prohibited. Fines for violations can be as stiff as 10% of a company's previous year's sales.

The 2008 antitrust law was somewhat vague, and businesses were confused as to whether they were in compliance or not. There were also very few investigations under the new law, and so precedents weren't available to guide behavior. The new rules are more specific, and their goal is to promote fair market competition.

It didn't take long for the first test case. In February 2011, Hudong.com (an online encyclopedia) claimed that Baidu (a search engine with a market share of 75.5%; Google is second with 19.6%—after

reducing its presence in China) blocks its web pages in favor of its own encyclopedia (Baidu Baike). Hudong claimed that Baidu had used its dominant position to "bully and block competitors." The Chinese investigative agency (State Administration for Industry and Commerce) may hear the case. Under the 2008 antitrust law, the three complaints filed against Baidu were dismissed.

Fang Xingdong, chairman of the board of Chinalabs.com (a Beijing IT think tank), states that Baidu, Tencent, and Alibaba are the leading oligopolies in China's Internet market, and that the government "urgently needs to formulate new rules to prevent the abuse of dominant position."

Source: Lan Lan, "New Regulations to Curb Price Monopolies," *China Daily*, January 5, 2011, at www.chinadaily.com.cn/bizchina/2011-01/5/content_11797083.htm; Wang Xing, "Baidu Accused of Abusing Dominant Position," *China Daily*, February 23, 2011, at www.chinadaily.com.cn/usa/business/2011-02/23/content_12063070.htm.

In Chapter 1 we cited the proposed 1997 merger of Staples with Office Depot. Tom Stemberg, the CEO of Staples and instigator of the proposed merger, felt that the combined entity would have a 4 to 6% share of the office supply market. The government disagreed. They defined the market as "category killers" (stores that could serve all of your office supplies needs) and felt the merger would lead to two or only one category killer in many markets and that such a result would be anti-competitive. Stemberg, on the other hand, defined the market as Wal-Mart, Dell on-line, Radio Shack, Seven Elevens, local stationery stores, and the like. Again, it is difficult to define market boundaries.

Antitrust policy will change between political administrations (based on how vigorously the executive branch pursues cases) and based on the constituency of the Supreme and other courts (the judicial branch of government). Although the legislative branch can produce new laws, it has not been active in this area; the enforcement and interpretation of existing laws determine antitrust policy today.

STRATEGY SESSION: Antitrust on the Continent

Reckitt Benckiser was fined £10.2 million by Britain's Office of Fair Trading for abusing its dominant position with respect to its popular heartburn medication Gaviscon. The firm's actions stopped National Health Services (NHS) doctors' ability to potentially prescribe cheaper alternative medications.

Here's what happened. When prescribing medication for a patient, NHS doctors can do a computer search for a branded medication. The search also gives them the names of cheaper generics that the doctor can prescribe instead. Reckitt circumvented this procedure by withdrawing and de-listing the original Gaviscon. Thus, when doctors searched for the medication, they wouldn't find it or its cheaper substitutes.

So how did Reckitt violate UK and European competition law? It used high pressure sales tactics on doctors to prescribe Gaviscon Advance, a slightly altered version of Gaviscon. Gaviscon Advance was listed in the NHS computer system, but there were no cheap generics listed. So doctors prescribed the expensive Gaviscon Advance. Because Reckitt admitted guilt and cooperated in the investigation, the fine was reduced from £12 million. The UK has a whistle-blower law that rewards whistle-blowers for

turning in corporate cheaters. A whistle-blower produced e-mails that purportedly showed that Reckitt personnel had tried to block a generic to Gaviscon Advance being developed. It was stated that the lack of a generic substitute had cost the NHS millions of pounds.

The chief executive, John Fingleton, of the Office of Fair Trading stated the importance of this case: "Vigorous competition between firms supplying the public sector is vital to ensure that taxpayers get the best value for money. The case underlines our determination to prevent companies with a dominant position in a market from using their strength to seek to restrict competition. The imposition of penalties should serve to deter firms from engaging in anti-competitive behavior of this sort in the future."

Sources: Julia Kollwe, "Reckitt Benckiser Fined £10.2m by OFT," *The Guardian*, Oct 15, 2010, www.guardian.co.uk/business/2010/oct15/reckitt-benckiser-fined-oft-gaviscon.

THE PATENT SYSTEM

While the antitrust laws are designed to limit monopoly, not all public policies have this effect. The patent system is a good example. U.S. patent laws have granted the inventor exclusive control over the use of an invention for 20 years (from initial filing), in exchange for his or her making the invention public knowledge.

Three principal arguments are used to justify the existence of the patent laws. First, these laws are regarded as an important incentive to induce inventors to put in the effort required to produce inventions. Particularly in the case of the individual inventor, it is claimed that patent protection is a strong incentive. Sec-

ond, patents are regarded as a necessary incentive to induce managers to carry out further work and make the necessary investments in pilot plants and other items required to bring inventions to commercial use. If an invention became public property when made, why should managers incur the costs and risks involved in experimenting with a new process or product? Managers at another firm could watch, take no risks, and duplicate the process or product if it is successful. Third, it is argued that because of the patent laws, inventions are disclosed earlier than otherwise, the consequence being that other inventions are facilitated by earlier dissemination of information.

Unlike most other goods, new technological knowledge cannot be used up. A person or firm can use an idea repeatedly without wearing it out, and the same idea can serve many users at the same time. No one need be getting less of an idea because others are using it, too. This fact creates an important difficulty for any firm that would like to make a business of producing knowledge. For an investment in research and development to be profitable, managers must be able to sell the results, directly or indirectly, for a price. But potential customers are unwilling to pay for a commodity that, once produced, becomes available to all in unlimited quantity. There is a tendency to let someone else pay for it if it then becomes available for nothing.

The patent laws, which are a way of addressing this issue, make it possible for managers to produce new knowledge and sell or use it profitably. But the patent system has the disadvantage that new knowledge is not used as widely as it should be because the patent holder, who attempts to make a profit, sets a price sufficiently high that some people who could make productive use of the patented item are discouraged from doing so. From society's point of view, all who can use an idea should be allowed to do so at a very low cost because the marginal cost of their doing so is often practically zero. However, this shortsighted policy would provide little incentive for invention.

Without question, the patent system enables innovators to appropriate a larger portion of the social benefits from their innovations; but this does not mean that patents are effective in this regard. Contrary to popular opinion, patent protection does not make market entry impossible or even unlikely. Within four years of their introduction, 60% of the patented successful innovations included in one study were imitated.⁸ Nonetheless, patent protection generally increases imitation costs. In that study, the median estimated increase in imitation cost (the cost of developing and commercially introducing an imitative product) was 11%. In the drug industry, patents had a bigger impact on imitation costs than in the other industries, which helps to account for the fact that patents are regarded as more important in drugs than elsewhere. The median increase in imitation cost was about 30% in drugs, in contrast to about 10% in chemicals and about 7% in electronics and machinery.

8. E. Mansfield, M. Schwartz, and S. Wagner, "Imitation Costs and Patents: An Empirical Study," *Economic Journal*, December 1981.

STRATEGY SESSION: Monopoly and Social Issues

A French doctor named Andree Briend has invented a food product that could save the world from hunger and malnutrition. It's called Plumpy'nut, and it's made from peanuts. Since most peanuts are grown in poor countries, and the rest of the ingredients are quite simple to acquire, assembly of Plumpy'nut on-site is quite easy (in fact, Briend once concocted a batch in a blender in Malawi to show how easy it could be made). Plumpy'nut is an edible paste chock-full of vitamins and calories. Miraculous stories of its use report children going from near death to relative health in a month. And it's not just a few cases; it works for the overwhelming majority of those who take it. In addition, the treatment is rendered at home and not in third-world hospitals, which are likely to be riddled with infectious diseases. Furthermore, it is taken without water (which is most likely contaminated in the countries where it's currently used).

Plumpy'nut is, however, a brand name registered trademark of Nutriset, a private French company that first produced and sold the product, and Nutriset's owner is listed as its co-inventor. Legally, Nutriset has all the rights of a patent owner. Nutriset realizes the

potential of its product and aggressively protects its intellectual property. The product is produced mostly in France (it has established franchisees in other countries), and UNICEF buys 90% of its Plumpy'nut from Nutriset in France. This suggests lots of transportation of peanuts and high production costs.

Patents and exclusive distribution rights exist to encourage innovation and to take advantage of scale economies. But some people question whether this protection of intellectual property should apply to all products, such as those that could benefit so many people (according to the United Nations, 925 million people, or 13.6% of the world's population, suffer from malnutrition). Because of Nutriset's monopoly power, the cost of a two-month treatment of Plumpy'nut costs \$60. As Patricia Wolff, a St. Louis pediatrician and founder of Med and Food for Kids, states, "Poverty is a business. There's money to be made and there are people who have that kind of way of thinking."

Nutriset is being sued by Breedlove Foods and the Mama Cares Foundation with the claim that it is using anticompetitive practices to protect its leading market share. As an indication of Nutriset's monopoly power

TRADE AND TRADE POLICY

To ask why we have a trade policy implies that we need to ask why we trade. Once we answer that question, we can use the same tools developed in earlier chapters to explain trade policy.

Foreign Trade

First, it is important to recognize that foreign trade is of great importance to the United States (and to all but the most isolated countries of the world). There are products our citizens desire but can't produce (such as kiwis or bananas); products we can produce that can be produced less expensively or with higher quality elsewhere (most consumer electronic equipment); products we produce that others can't (jetliners for South Africa); and products we can produce less expensively or with higher quality (precision machine tools for Panama). As Table 17.2

to restrict the use of the product, some estimate only 10–15% of those who would benefit from Plumpy'nut receive it. The implication is that increased production close to the consumption site would lower the cost and increase use. The crux of the issue is that the market is viewed not only as a cure for those afflicted by malnutrition but also as a preventive measure so that children don't become malnourished in the first place. This raises the possibility of market with current sales of \$300 million annually going to a \$6 billion a year market. That's a market worth protecting (or so the plaintiffs in the lawsuit imply). An additional problem is that Plumpy'nut's patent is very broad and doesn't expire until 2017. Nutriset is concerned that one or more of the world's large food producers, such as Pepsico, might get in the market.

Meanwhile, the clock is ticking. There's a "golden interval" before age two during which it is especially important to combat malnutrition. During the first two years of life, brain growth and development of the body's immune system are extensive. If malnutrition occurs in those years, its stunting impact on brain and immune system development is

generally irreversible, making it difficult for people to get out of poverty (malnutrition and poverty usually go hand-in-hand). Others argue that freeing up the production would produce local jobs for farmers and manufacturers.

The question then is whether some patentable items are just too critical to human welfare to be patented. If such a judgment were to be made, would there be a bright line between what is and what isn't patentable. And would such a possibility create a chilling effect on experimentation? Briend's story is interesting in this respect. He knew nothing about food manufacturing and marketing. He signed a consulting agreement with Nutriset, and Nutriset's brought the product to market. Briend, legend has it, has not taken any profits from his creation, so perhaps development without patent protection is possible.

Source: Andrew Rice, "The Peanut Solution," *New York Times*, September 2, 2010, at www.nytimes.com/2010/09/05/magazine/05Plumpy-t.html.

shows, we exported about \$1.84 trillion in goods and services in 2008 and exported \$1.57 trillion in goods and services in 2009 (showing the dramatic impact of the global recession). Goods accounted for a little over a trillion dollars in 2009, with 36.5% of this in capital goods (except automotive) and 27.8% in industrial supplies and materials. We imported almost \$2.54 trillion in goods and services in 2008 and imported \$1.93 trillion in goods and services in 2009 (again showing the recession's impact). About \$1.57 trillion in 2009 paid for imports of goods, much of them capital goods (except automotive), consumer goods (nonfood, except automotive), industrial supplies and materials, and automotive vehicles, engines, and parts. As Table 17.3 shows, about 32% of exported goods went to our nearest neighbors (Canada and Mexico) in 2010, and slightly under 52% of our imports of goods were from Canada, China, Mexico, and Japan, with China's share about 1.3 times larger than Canada's. Fifteen nations accounted for about 72% of

STRATEGY SESSION: Using Patents to Maintain Market Power

The U.S. government often tries to break up monopolies or to limit their market power, but one government institution, the patent system, is designed to grant firms monopoly power over their inventions and product innovations. The government grants companies patents for 20 years, giving the patent holder exclusive rights to an invention or new product for the life of the patent. Because companies apply for a patent before it is on the market, the effective life of a patent can be significantly less than 20 years. For drugs (a major user of patents), the effective life of a patent is between 7 and 12 years.^a

Thanks to the monopoly granted by the patent system (and the anticipated monopoly profit from a popular product), drug companies are willing to invest significant funds in research and development (R&D) in the hopes of discovering a blockbuster drug. There have been estimates that a successful new drug can cost between \$0.8 billion to \$1.7 billion. These estimates have been criticized for what they include as costs, but there is no doubt that the amount spent is substantial.^b Without patent protection, it is doubtful that drug companies would undertake highly problematic and expensive R&D if another firm could just buy the drug, reverse-engineer it to determine its composition, and then manufacture it as a generic without having the R&D expense.

What do managers of drug companies do price-wise with their patent-generated monopoly power? They make hay while the sun shines. When the patent runs out in 7 to 12 years, other manufacturers can reverse-engineer the pill and produce it on the cheap. One test is to compare the price of drugs as the drugs come off patent with the price of generics that then appear in the market.

Arimidex, a commonly prescribed breast cancer drug, was released as a generic (Anastrozol) in June 2010. Costco, who typically marks up the wholesale

price by 14–15%, sells a month's supply of Anastrozol for \$27. A month's supply of Arimidex sells for over \$400. That's a 15 times markup of the patented drug over the generic—a hefty markup (and it shows the ability of a monopolist to control price).

But the story is worse than it initially seems for Arimidex users (unless they are members of Costco). Major drugstore chains and some other big-box stores with pharmacies were selling the new generic at about 13 times its price at Costco.^c This is attributed to the fact that customers were used to paying the high price for Arimidex, and so a lower relative price on Anastrozol would seem like a bargain. The real bargain is at Costco (and the price of the generic is likely to come down as the word of Costco's price spreads).

The hero in the story of the drug companies' markups are TV stations' investigative reporters.^d Starting in Detroit in 2002 and later in other cities, such as Philadelphia, reporters would compare the patented drug with the generic and report, of course, on the ones with 1,000% markups. But remember the trade-off: Would the drug companies do any R&D if they couldn't recoup the costs of all the failed drugs?

^a4th Annual Conference on Drug Discovery and Therapy at www.icddt.com/TrackDescription/CPCO-track-desc.htm.

^bRoger Collier, "Drug Development Costs Hard to Swallow," at www.ncbi.nlm.nih.gov/pmc/articles/PMC2630351/.

^c"Arimidex Price Markup at Major Pharmacy Chains" at www.kevinmd.com/blog/2010/10/arimidex-price-markup-major-pharmacy-chains.html.

^dSee www.truthorfiction.com/rumors/g/generics.htm for story on Detroit, Florida and Philadelphia.

TABLE 17.2**U.S. Goods and Services by General Type Exports and Imports: 2008 and 2009 in Billions of Dollars**

Product	Exports		Imports		Trade Surplus	
	2008	2009	2008	2009	2008	2009
Food, Feeds & Beverages	108.3	93.9	89.1	81.6	19.2	12.3
Industrial Supplies & Materials	388.0	296.7	779.5	462.5	-391.5	-165.8
Capital Goods Except Automotive	457.7	390.5	453.7	369.3	4	21.2
Automotive Vehicles, Parts & Engines	121.5	81.7	231.2	157.6	-109.7	-75.9
Consumer Goods (not Food or Auto)	161.3	150.0	481.6	428.4	-320.3	-278.4
Total Goods	1,304.9	1,068.5	2,139.6	1,559.6	-834.7	-491.1
Services	534.1	502.3	398.3	370.3	135.8	132
Total Goods & Services	1,839.0	1,570.8	2,537.9	1,929.9	-698.9	-359.1

Source: Calculated from the Economic Report of the President, February 2011, Table B-106.

TABLE 17.3**U.S. Total Trade in Goods (Billions of Dollars) and Top Trading Partners, 2010**

	Exports	Imports	Total	Percentage of Total Trade	Trade Deficit	Percentage of Deficit
Total All Countries	1,278.1	1,912.1	3,190.2	100.0	634.0	100.0
Top 15 Countries	894.1	1,401.3	2,295.4	72.0	507.2	80.0
Canada	248.8	276.5	525.3	16.5	27.7	4.4
China	91.9	364.9	456.8	14.3	273.0	43.1
Mexico	163.3	229.7	393.0	12.3	66.4	10.5
Japan	60.5	120.3	180.8	5.7	59.8	9.4
Federal Republic of Germany	48.2	82.7	130.9	4.1	34.5	5.4
United Kingdom	48.5	49.8	98.3	3.1	1.3	0.2
South Korea	38.8	48.9	87.7	2.7	10.1	1.6
France	27.0	38.6	65.6	2.1	11.6	1.8
Taiwan	26.0	35.9	61.9	1.9	9.9	1.6
Brazil	35.4	23.9	59.3	1.9	-11.5	-
Netherlands	35.0	19.0	54.0	1.7	-16.0	-
India	19.2	29.5	48.7	1.5	10.3	1.6
Singapore	29.1	17.5	46.6	1.5	-11.6	-
Venezuela	10.7	32.8	43.5	1.4	22.1	3.5
Saudi Arabia	11.6	31.4	43.0	1.3	19.8	3.1

Source: Calculated from www.census.gov/foreigntradestatistics/highlights/top/top1012yr.html.

PROBLEM SOLVED: Government Purchase of Toxic Assets

Through its actions, the government can also increase the efficiency of markets. This is particularly useful when markets fail because, as discussed later on page 716, the Treasury proposed to purchase some CDOs (Collateralized Debt Obligations) and MBSs (Mortgage-Backed Securities) from financial institutions. We show the possible outcomes of such a plan using a simplified situation.

Suppose there are only two types of risky bonds: Merely illiquid (*I*) types and truly toxic (*T*) types. Each type of bond is worth \$0 if it defaults, and \$1,000 if it does not default. Type *I* bonds have a 10% chance of default, and type *T* bonds have a 30% chance of default.

Assume there are only two current bondholders, each holding a different type of bond. Each current bondholder knows the type of bond they hold and is risk-averse with utility function $U = (W)^{0.5}$ where W is the bondholder's wealth. Imagine each bondholder's only wealth is its bond.

The sole buyer of these risky bonds is the U.S. Treasury. However, the Treasury does not know whether a bond is the illiquid type (*I*) or the toxic type (*T*) before buying it. Suppose the Treasury offers to buy any bond at a price of 700. Can we determine who

will sell to the Treasury and what will be the Treasury's expected profit? We can do so by using our expected utility analysis.

Suppose the government offers 700 for the bond. Then a seller will have expected utility of

$$EU_{\text{SellBond}} = 700^{0.5} = 26.458$$

If the type *T* holder retains the bond, her expected utility is

$$EU_{\text{RetainBondT}} = 0.7 \cdot 1,000^{0.5} + 0.3 \cdot 0^{0.5} = 22.136$$

Since $26.458 > 22.136$, the type *T* seller will sell the bond to the government.

If the type *I* holder retains the bond, her expected utility is

$$EU_{\text{RetainBondI}} = 0.9 \cdot 1,000^{0.5} + 0.1 \cdot 0^{0.5} = 28.461$$

Since $28.461 > 26.458$, the type *I* seller will retain her bond.

So under this scheme, the Treasury will only purchase toxic bonds. Holders of type *I* bonds choose not to sell. If the government pays 700 and gets the type *T* bond (whose expected value is 700, i.e.,

all U.S. trade in goods and 80% of our trade deficit. Of our 15 largest trading partners, we ran a trade surplus in goods only with Brazil, the Netherlands, and Singapore. The largest trade deficits (import value minus export value) were with China (\$273 billion), Mexico (\$66.4 billion), and Japan (\$59.8 billion). Also of interest is the U.S. trade deficit with the two OPEC countries represented (Venezuela and Saudi Arabia). Overall, we had a goods trade deficit of \$634 billion in 2010. We traditionally run a trade surplus in services (and it was \$132 billion in 2009). Our largest trade deficit was in 2006 at \$839.5 billion. Way back in 1952, we ran a trade surplus of \$2.6 billion.

Why does trade occur among countries? As economists have pointed out for over two centuries, trade permits specialization, and specialization increases output. Because the United States can trade with other countries, it can specialize in

$0.7*1,000 + 0.3*0$], the government breaks even, i.e., $700 - 700 = 0$.

Now, suppose that the Treasury changes its offer to

1. Purchase any bond for the price of 700.
or
2. Purchase 25% of a bond for a price of 220. (In this case, the current bondholder keeps the remaining 75% stake in the bond.)

Each current bondholder can choose whether to accept offer 1 or 2. In this case, can we determine who will sell to the Treasury and what will be the government's profit? Again we can use our expected utility analysis. If either seller takes offer 1, her utility is

$$700^{0.5} = 26.458.$$

Suppose seller T takes offer 2. Her utility is

$$\begin{aligned} EU_{Option2T} &= 0.7*(220 + 0.75*1,000)^{0.5} + 0.3*[220]^{0.5} \\ &= 0.7*(970)^{0.5} + 0.3*[220]^{0.5} = 26.251 \end{aligned}$$

If she holds the bond, her expected utility is 22.136, as it was under the previous plan. Since $26.458 >$

$26.251 > 22.136$, she will sell the bond to the government under offer 1.

Suppose seller I takes offer 2. Her utility is

$$\begin{aligned} EU_{Option2I} &= 0.9*(220 + 0.75*1,000)^{0.5} + 0.1*[220]^{0.5} \\ &= 0.9*(970)^{0.5} + 0.1*[220]^{0.5} = 29.514 \end{aligned}$$

If she holds the bond, her expected utility is 28.461 as it was under the previous plan. Since $29.514 > 28.461 > 26.458$, she will sell the bond to the government under offer 2.

Thus, the government's expected revenue is

$$\begin{aligned} &0.7*1,000 + 0.3*0 + 0.25[0.9*1,000 + 0.1*0] \\ &= 700 + 0.25*900 = 700 + 225 = 925 \end{aligned}$$

and the government's expense is

$$700 + 220 = 920 \text{ yielding a profit of } 925 - 920 = 5$$

So, it is possible to construct a scheme where the Treasury can buy back bad bonds, increase the well-being of their holders (i.e., increase the holders' utilities), and make a profit in the process. Notice that the government accomplishes this by shouldering some of the risk of merely illiquid types and all of the risk of the toxic types.

the goods and services it produces well and cheaply. Then it can trade them for goods and services that other countries are particularly good at producing. The result is that we and our trading partners benefit (as we will demonstrate next).

International differences in resource endowments and the relative quantity of various types of human and nonhuman resources are important bases for specialization. Consider countries with lots of fertile soil, little capital, and much unskilled labor. They are likely to find it advantageous to produce agricultural products, whereas countries with poor soil, much capital, and highly skilled labor probably do better to produce capital-intensive, high-technology goods. However, the basis for specialization does not remain fixed over time. Instead, as technology and resource endowments of various nations change, the pattern of international specialization changes as well. For example, the United States specialized more in

STRATEGY SESSION: Making Whistle-Blowing Pay Off

Want to pick up a quick \$195,000? The British Office of Fair Trading (the equivalent of the U.S. Department of Justice's Antitrust Division) is offering a £100,000 payment on its website for whistle-blowers whose information leads to the apprehension and conviction of price fixing conspirators. A hotline telephone number is given. The chief executive of the office, John Fingleton, calls the practices he has been going after "shoddy, complacent, and cozy."

He has prosecuted supermarkets for their pricing practices on such items as toothpaste and tea. He is in the midst of his biggest investigation to date: 112 construction companies were accused of colluding to rig bids for construction projects involving hospitals, schools, and universities. Seventy-seven companies have already admitted wrongdoing. The conspiracy involved sharing bid information, predetermining which company would win the bid, and then having the other companies submit higher bids than the "winner." Fingleton is also investigating the pricing of cigarettes and marine hoses used in the oil

industry. In 2007 he registered record fines, including a £121.5 million judgment against British Air for fuel surcharge fixing and £116 million from supermarkets and dairies for fixing the price of milk.

Why this new interest in price fixing? The trend is global (see the other Strategy Sessions describing U.S. and EU enforcement). Part of it is attributable to pressure from the United States, and increasingly from the European Union, for an international crackdown on price fixing. The European Union has imposed fines of over \$9.1 billion since 2005 (through the first third of 2008), which is more than it levied in the previous 15 years. South Korea, Japan, India, and Brazil are starting to pay attention, as is Hong Kong. This is attributed by a U.S. deputy assistant attorney general to a "growing worldwide consensus that international cartel activity is pervasive and is victimizing businesses and consumers everywhere."

Fingleton describes the situation as "a big change in Europe. And it's quite sudden." Here is a table showing the largest recent U.S. and EU antitrust fines.

raw materials and foodstuffs a century ago than it does now. India and China have the same story, but only 20 years ago they were more agrarian. (For example, the U.S. trade deficit with China in 1985 was only \$6 million.)

Using Demand and Supply to Determine the Country of Import and the Country of Export

How can a manager predict whether his or her country has a comparative advantage in the production of a particular product? One important indicator is whether the country's firms can make money by producing and exporting the product. Consider the Wilton Company, the maker of a new product produced in the Netherlands and the United States—the only two countries where this product has a significant market. In the United States, the demand curve for the product is such that

$$Q_D^U = 8 - P_U \quad (17.5)$$

United States			
Company	Country	Year	Fine (Million \$)
Hoffman-LaRoche	Switzerland	1999	500
Korean Air Lines	S. Korea	2007	300
British Airways	United Kingdom	2007	300
Samsung	S. Korea	2006	300
BASF	Germany	1999	225
Hynix Semiconductor	S. Korea	2005	185
Infineon Technologies	Germany	2004	160
SGL Carbon	Germany	1999	135
Mitsubishi Corp	Japan	2001	134
Ucar International	United States	1998	110

European Union			
Company	Country	Year	Fine (Million £)
ThyssenKrupp	Germany	2007	480
Hoffman-LaRoche	Switzerland	2001	462
Siemens	Germany	2007	397
Eni	Italy	2006	272
Lafarge	France	2002	250
BASF	Germany	2001	237
Otis	United States	2007	225
Heineken	Netherlands	2007	219
Arkema	France	2006	219
Solvay	Belgium	2006	167

Two interesting observations emerge from the table. The average U.S. case occurred midway through 2002, whereas the average EU case occurred in 2005 (reflecting more recent activity in the European Union than in the United States); and the average value of the fines in the United States for 2007 was \$300 million, whereas the average fine in the European Union in 2007 was \$434 million.

The anti-corruption activity continues unabated in England. The government's National Fraud Reporting Centre has opened a hotline to allow all cases to be handled centrally. They have landmark cases in the use of civil recovery powers (against Balfour Beatty) and the first prosecution of a British executive (Robert Dougall of DePuy International) in a bribery case. They also matched a record fine against Aon Limited.

Source: Michael Peel, "Tough Protection: Competition Authorities Are Clamping Down," *Financial Times*, May 8, 2008, p. 9; and "Whistleblowing: A New UK Helpline," at www.nortonrose.com/expertise/business-ethics-and-anti-corruption.

and the supply curve is such that

$$Q_S^U = -2 + P_U \quad (17.6)$$

where P_U is the price of a unit of the product (in dollars) in the United States (and $P_U \geq 2$), Q_D^U is the quantity demanded (in millions of units) per month in the United States, and Q_S^U is the quantity supplied (in millions of units) per month in the United States.

In the Netherlands, the demand curve for this product is such that

$$Q_D^N = 6 - 2P_N \quad (17.7)$$

and the supply curve is such that

$$Q_S^N = -2 + 2P_N \quad (17.8)$$

where P_N is the price of a unit of the product (in euros) in the Netherlands (and $P_N \geq 1$), Q_D^N is the quantity demanded (in millions of units) per month in the

Netherlands, and Q_S^N is the quantity supplied (in millions of units) per month in the Netherlands.

Because the new product is being introduced for the first time in the Netherlands and the United States, managers and analysts in both countries would like to predict whether, after markets in both countries settle down, this product will be exported and, if so, by which of these two countries. To answer this question, we must begin by noting that if the cost of transporting this product from the United States to the Netherlands (or vice versa) is zero (which we will assume for simplicity), the price of this product after trade must be the same in both countries. Why? Because if it were different, a firm could make money by purchasing it in the country where its price is lower and selling it in the country where its price is higher. As this continues, the price would rise in the former country and fall in the latter country, until eventually the price in both countries would be equal.

What do we mean by the prices in both countries being equal? Prices in the United States are quoted in dollars; prices in the Netherlands are quoted in euros. What we mean is that based on prevailing exchange rates, the prices in both countries are the same. If the U.S. dollar exchanges (at banks and elsewhere) for \$2 per euro, a price of \$10 in the United States is equivalent to a price of 5 euros in the Netherlands. Consequently, if this is the exchange rate, the prices in the two countries being the same means that

$$0.5P_U = P_N \quad (17.9)$$

If there is no government intervention in the market for this product and the market is competitive, the price of this product will tend to be at the level where the total world demand for the product equals the total world supply. In other words, in equilibrium

$$Q_D^U + Q_D^N = Q_S^U + Q_S^N \quad (17.10)$$

Using equations (17.5) to (17.8), we can express each of the values of Q in equation (17.10) as a function of P_U or P_N . Substituting each of these functions for each of the Q values in equation (17.10), we obtain

$$(8 - P_U) + (6 - 2P_N) = (-2 + P_U) + (-2 + 2P_N)$$

Substituting $0.5 P_U$ for P_N , we find that

$$\begin{aligned} (8 - P_U) + [6 - 2(0.5P_U)] &= (-2 + P_U) + [-2 + 2(0.5P_U)] \\ 14 - 2P_U &= -4 + 2P_U \\ 18 &= 4P_U \\ P_U &= \$4.5 \end{aligned}$$

Because $0.5P_U = P_N$, $P_N = 0.5(4.5) = 2.25$ euros. In other words, the price of the product is \$4.50 in the United States and 2.25 euros in the Netherlands.

Given these prices, we can determine whether the United States or the Netherlands will export the product. Based on equation (17.5), the monthly quantity demanded in the United States will be $8 - P_U = 8 - 4.5 = 3.5$ million units. Using equation (17.6), the quantity supplied per month in the United States will be $-2 + P_U = -2 + 4.5 = 2.5$ million units. Therefore, the United States will import $3.5 - 2.5 = 1$ million units per month. Based on equation (17.7), the monthly quantity demanded in the Netherlands will be $6 - 2P_N = 6 - 2(2.25) = 1.5$ million units. Based on equation (17.8), the quantity supplied per month in the Netherlands will be $-2 + 2P_N = -2 + 2(2.25) = 2.5$ million units. Therefore, the Netherlands will export $2.5 - 1.5 = 1$ million units per month.

In sum, the Netherlands will be the exporter of this new product, and its exports will equal 1 million units per month.

Analyzing the Argument for the Government's Advocacy of Free Trade Using Producer and Consumer Surplus

We now view the gains from free trade for both the United States and the Netherlands in the preceding example. To do so, we use the concepts of consumer surplus and producer surplus. Figure 17.3 shows the situation in the United States.

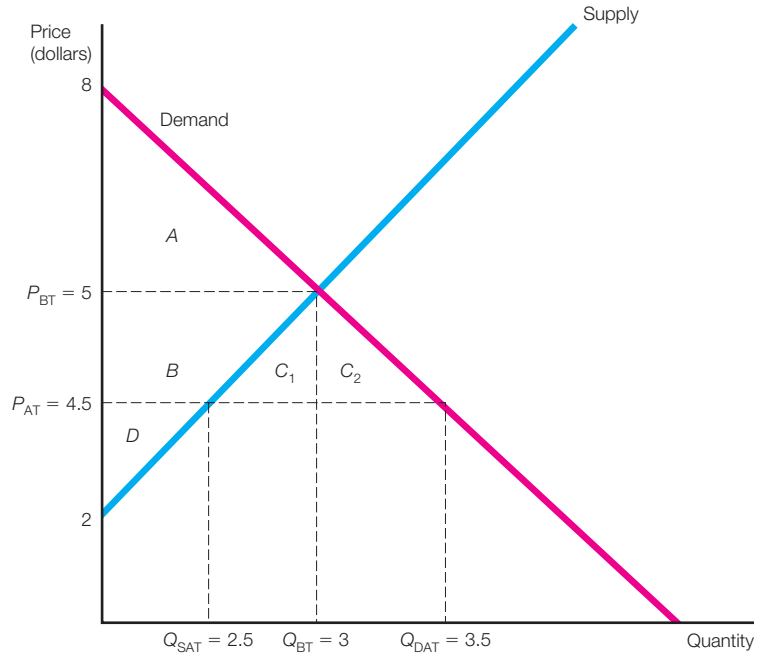
Before trade with the Netherlands, the price of the product in the United States was \$5, and 3 million units were transacted (calculated by setting equation (17.5) equal to equation (17.6) and solving for P_U). This left U.S. consumers with a consumer surplus of $A (= 0.5 \times 3 \times 3 = 4.5)$ and U.S. producers with a producer surplus of $B + D (= 0.5 \times 3 \times 3 = 4.5)$. Because of the lower price (\$4.50) after trade, consumers experience an increase in consumer surplus (now $A + B + C_1 + C_2 = 0.5 \times 3.5 \times 3.5 = 6.125$). Thus consumers gain $B + C_1 + C_2 (= 6.125 - 4.5 = 1.625)$. Prior to trade, U.S. producers received a producer surplus of $B + D$, which has now been lowered to $D (= 0.5 \times 2.5 \times 2.5 = 3.125)$ as a result of the lower price. Therefore U.S. producers lose $B (= 4.5 - 3.125 = 1.375)$ in producer surplus because of trade. The sum of the consumers' gain ($B + C_1 + C_2$) and the producers' loss (B)—that is, $C_1 + C_2$ —is the gains from trade for the United States. Its magnitude is $0.5 \times 0.5 \times 1 = 0.25 = 1.625 - 1.375$. Society is better off because social welfare has increased from $A + B + D$ to $A + B + D + C_1 + C_2$, or by $C_1 + C_2$. Although the producers have lost, in theory there are enough gains to consumers that producers could be more than compensated for their loss.

For example, suppose the U.S. government wrote the producers a check for $B + C_1$. Producers now have gained C_1 (that is, $B + D - B + B + C_1 = B + D + C_1$ versus $B + D$ before the trade). Where did the government get the money to pay the producers? By taxing the recently better-off consumers. However, even after being taxed $B + C_1$, consumers are still better off because they have $A + B + C_1 + C_2 - (B + C_1) = A + C_2$, as compared to the A they had

FIGURE 17.3

Consumer and Producer Surplus in the United States Before and After Trade

The U.S. gain in consumer surplus is $B + C_1 + C_2$; the U.S. loss in producer surplus is B , for a net societal gain of $C_1 + C_2$ as the result of free trade. P_{AT} = price after trade; P_{BT} = price before trade; Q_{SAT} = quantity supplied after trade; Q_{DAT} = quantity demanded after trade; and Q_{BT} = quantity demanded and supplied before trade.



before trade. Thus trade makes the United States better off and, depending on how the gains are distributed, *can* make both consumers and producers better off. One role of government is to make decisions about redistribution.

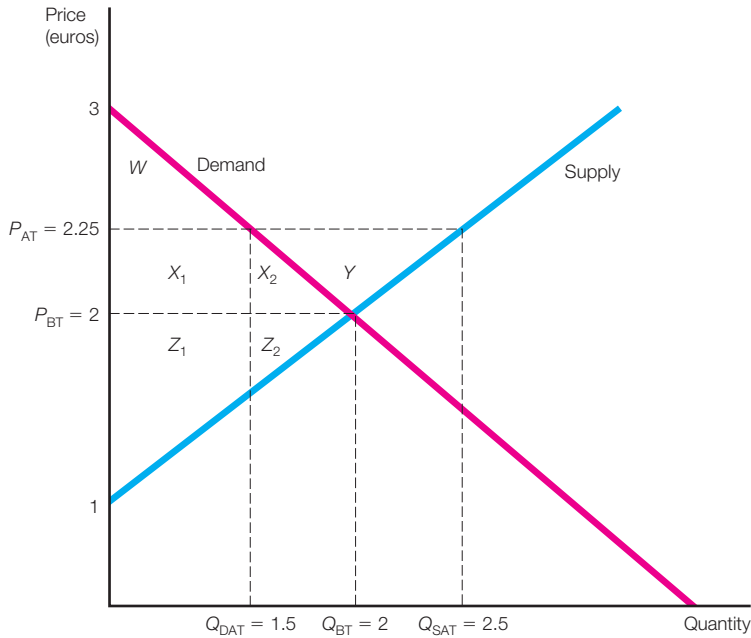
Figure 17.4 shows the situation in the Netherlands. Before trade the good sold for 2 euros, and 2 million units were transacted (calculated by setting equation (17.7) equal to equation (17.8) and solving for P_N). Consumers in the Netherlands enjoyed a consumer surplus of $W + X_1 + X_2 (= 0.5 \times 1 \times 2 = 1)$, while the Dutch producers had a producer surplus of $Z_1 + Z_2 (= 0.5 \times 1 \times 2 = 1)$ for a total Dutch social welfare of $W + X_1 + X_2 + Z_1 + Z_2 (= 2)$.

After trade, the price increases to 2.25 euros. The price increase hurts Dutch consumers, and their consumer surplus falls to $W (= 0.5 \times 0.75 \times 1.5 = 0.5625)$ —a loss of $X_1 + X_2 (= 1 - 0.5625 = 0.4375)$. However, the price increase benefits Dutch producers, and the producer surplus increases to $X_1 + X_2 + Y + Z_1 + Z_2$

FIGURE 17.4

Consumer and Producer Surplus in the Netherlands Before and After Trade

The Dutch gain in producer surplus is $X_1 + X_2 + Y$, and the Dutch loss in consumer surplus is $X_1 + X_2$, for a net societal gain Y as the result of free trade. P_{AT} = price after trade; P_{BT} = price before trade; Q_{SAT} = quantity supplied after trade; Q_{DAT} = quantity demanded after trade; and Q_{BT} = quantity demanded and supplied before trade.



($= 0.5 \times 1.25 \times 2.5 = 1.5625$), a gain of $X_1 + X_2 + Y$ ($= 1.5625 - 1 = 0.5625$). The producer gain more than offsets the consumer loss (by $Y = 0.5 \times 0.25 \times 1 = 0.5625 - 0.4375 = 0.125$). The social welfare is now $W + X_1 + X_2 + Y + Z_1 + Z_2$ (up by Y). Although the gains in each country are equal (that is, 0.125 euros in the Netherlands equals 0.25 dollars in the United States), this does not have to be the case. The point is that both countries have gained from the trade (by $C_1 + C_2$ in the United States and by Y in the Netherlands), giving each country the incentive to engage in trade.

One thing, however, must be true in this two-country, one-good trading world: The United States imports must equal the Dutch exports. The 1 million units imported are the U.S. physical trade deficit. The fiscal trade deficit is the 1 million units multiplied by the world price of \$4.50, or \$4,500,000. The Netherlands, on the other hand, is running a physical trade surplus of 1 million units and a fiscal trade surplus of 2.25 million euros (or \$4.5 million).

How does the United States obtain the euros to import these goods from the Netherlands? We can receive euros from the Netherlands from Dutch imports of other American products; from Dutch tourists who visit the United States; and from Dutch citizens or businesses or governments that buy U.S. securities, services, real estate, and so on. In addition, we may have reserves of euros that were accrued in the past.

Use of Tariffs and Quotas to Mitigate the Gains from Trade

Although our example showed that the United States gained from trade with the Netherlands, without any redistribution of gains, the consumers gain and the producers lose. Suppose increased taxes on consumers are unpopular, so business will not receive any share of the consumer gains. There's a large loss for producers (B) and not many producers. If there are only a few producers (n) of equal size, their share of the loss is B/n —so each has a lot to lose from trade. They may have a trade association that can persuade Congress to pass legislation to curb imports of the product from the Netherlands either by placing a tariff (T) on the product (hence raising the price of the product in the United States from P_{AT} to $P_{AT} + T$ and preserving some of the before-trade producer surplus of $B + D$) or by placing a quota on imports (so that the imports allowed into the United States would be only a fraction of the $Q_{DAT} - Q_{SAT}$ imported under free trade). They may be willing to spend significant money (up to B/n per firm) to eliminate free trade.

Let's see what the impact of a quota of $\alpha(Q_{DAT} - Q_{SAT}) = Q_{DAQ} - Q_{SAQ}$ where $1 \geq \alpha > 0$ would be. This is shown in Figure 17.5.

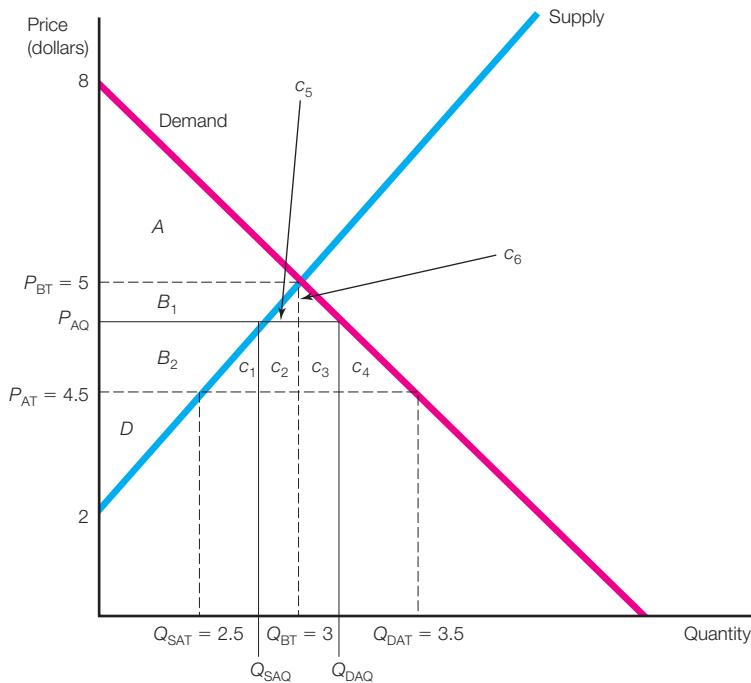
The quota raises the price in the United States to P_{AQ} . This decreases consumer surplus by $B_2 + c_1 + c_2 + c_3 + c_4$ and increases producer surplus by B_2 . Instead of the social welfare under free trade of $A + B_1 + B_2 + c_1 + c_2 + c_3 + c_4 + c_5 + c_6 + D$, we now have social welfare of $A + B_1 + B_2 + c_5 + c_6 + D$ in the United States. Thus we have a loss in social welfare in the United States of $c_1 + c_2 + c_3 + c_4$, and the gains from trade have shrunk to $c_5 + c_6$. The Dutch haven't lost all their gains from free trade because of the quota. They still export $Q_{DAQ} - Q_{SAQ}$ to the United States and receive a price premium of $P_{AQ} - P_{AT}$ for each unit. Thus they capture additional producer surplus of $c_2 + c_3$ for these decreased exports. The Dutch will suffer a loss in welfare, but it isn't as drastic as it could be because of the premium of $P_{AQ} - P_{AT}$ received for the quota items. With a quota, part of our loss in welfare is transferred to the exporting country as a gain to them.

Can the U.S. government accomplish the same objective of limiting the number of imports to $Q_{DAQ} - Q_{SAQ}$ with a smaller negative effect on social welfare? Suppose the government puts a tariff of $P_{AQ} - P_{AT} = T$ on each unit of the imported product. Then the price U.S. consumers pay for the product will be P_{AQ} , U.S. demanders will demand Q_{DAQ} units, U.S. suppliers will supply Q_{SAQ} units,

FIGURE 17.5

Consumer and Producer Surplus in the United States Before and After Trade with an Import Quota of $Q_{DAQ} - Q_{SAQ}$ Units

The U.S. gain in consumer surplus is $B_1 + c_5 + c_6$, and the U.S. loss in producer surplus is B_1 , for a net societal gain of $c_5 + c_6$, with an import quota of $Q_{DAQ} - Q_{SAQ}$ units. P_{AT} = price after trade; P_{BT} = price before trade; Q_{SAT} = quantity supplied after trade; Q_{DAT} = quantity demanded after trade; Q_{BT} = quantity demanded and supplied before trade; Q_{SAQ} = quantity supplied after quota; Q_{DAQ} = quantity demanded after quota; and P_{AQ} = price in the United States after the import quota of $Q_{DAQ} - Q_{SAQ}$ is imposed.



and $Q_{DAQ} - Q_{SAQ}$ will be imported from the Netherlands. U.S. social welfare will increase by $c_2 + c_3$ because now the U.S. Treasury will collect $T(Q_{DAQ} - Q_{SAQ})$ in revenues from the imported goods instead of letting that amount accrue to producers from the Netherlands (as with the quota). The domestic social welfare loss due to trade restrictions shrinks to $c_1 + c_4$. However, free trade would allow us to capture all of the previous $C_1 + C_2$ instead of just part of it ($c_5 + c_6$ under a quota or $c_2 + c_3 + c_5 + c_6$ under a tariff). Our conclusion is that if government wants to restrict trade, a tariff is a more efficient way (for domestic social welfare) to do so.

Trade Policy When the Market Is Not Perfectly Competitive

The previous analysis assumes that markets are perfectly competitive; they are not, as the bulk of this book has pointed out. We have suggested strategies for managers to optimize these noncompetitive situations. Likewise, there are strategies for governments to use in situations where trade involves imperfectly competitive situations.

Traditionally, economists have tended to argue that free trade is the best policy to promote the interests of society as a whole. They generally applauded the lowering of tariffs in the 1960s and 1970s and looked with disfavor on the growth of protectionism in the early 1980s. They again applauded the formation of NAFTA in the late 1980s and other free trade areas around the world. But some economists have begun to dispute these traditional beliefs. In their view, the U.S. government should control the access of foreign firms to our domestic markets, as well as promote the activities of our firms in foreign markets. For example, if particular high-technology industries generate large technological benefits for other domestic industries, the government may be justified in using subsidies or tariffs to protect and promote these industries. And if economies of scale mean that only two highly profitable producers can exist in the world market, the government may be justified in using subsidies or tariffs to increase the chances that a U.S. firm is one of the lucky pair.

According to these economists, there are strategic industries that, from the point of view of a particular country, are worth protecting in this way. However, it is difficult to identify which industries fall into this category and to estimate how much the country would gain from such policies. Consequently critics of such strategic trade policies worry that special interest groups can use such policies to advance their own interests, and not those of the nation as a whole. Given the vague criteria for identifying which industries should be protected, many industries can use these ideas to justify protection for themselves and their allies, regardless of whether it is merited.

That being said, let's view the use of strategic trade policy in action. There is an ongoing dispute between the U.S. government and the European Union regarding charges and countercharges of unfair government subsidization of Boeing (by the United States) and Airbus (by the European Union) in the development of commercial jet airliners. A comparable battle is being waged by Canada (Bombardier) and Brazil (Embraer) over government subsidization of regional jet aircraft.

To illustrate strategic trade policy, we can use game theoretic models. Suppose only two firms, Boeing and Airbus, are capable of producing a new 250-seat aircraft. Managers at each firm must decide whether to produce and market such a plane. Because Boeing has a head start, it can make this decision first. Figure 17.6 shows the payoff matrix for both firms. If either firm is the sole producer of the plane, it will make \$120 billion; but if both firms decide to produce and market such a plane, both will lose \$8 billion. Clearly Boeing managers, who have the first

FIGURE 17.6

Payoff Matrix: Airbus and Boeing

		Boeing	
		Produce new plane	Do not produce new plane
Airbus	Produce new plane	−\$8 billion, −\$8 billion	\$120 billion, \$0
	Do not produce new plane	\$0, \$120 billion	\$0, \$0

FIGURE 17.7

New Payoff Matrix: Airbus and Boeing

		Boeing	
		Produce new plane	Do not produce new plane
Airbus	Produce new plane	\$2 billion, −\$8 billion	\$130 billion, \$0
	Do not produce new plane	\$0, \$120 billion	\$0, \$0

move in this game, will decide to produce the plane; and Airbus managers, once they realize that Boeing is committed to this course of action, will decide not to produce it.

But Boeing is a U.S. firm, whereas Airbus is 67% owned by EU firms as a joint venture of French, British, German, and Spanish aerospace firms (with their governments' blessings and participation). If these European governments decide to pay Airbus a subsidy of \$10 billion *if and only if* it produces the plane, the game has quite a different outcome. The new payoff matrix is shown in Figure 17.7, and it is clear that Airbus managers will produce the plane regardless of whether Boeing commits itself to producing it. In other words, Airbus managers now have a dominant strategy as a result of the subsidy to produce the plane. And Boeing managers, recognizing that this will be the Airbus decision, will not find it profitable to produce it. Instead they will decide against production of the plane.

In effect, the European governments have taken the profit of \$120 billion away from Boeing and bestowed it on Airbus. Admittedly, they have to pay a subsidy of \$10 billion, but this is relatively small for assuring a profit of \$120 billion for the “home team.” Because the \$120 billion is profit, one could envision a subsequent \$10 billion tax to compensate for the government subsidy. But the story gets even better. Since the EU government owns 67% of the aerospace firms, they got 67% of the \$120 billion profit (= \$80 billion). This is quite a handsome return on an investment of \$10 billion. Without question, this example seems to indicate that government intervention of this sort can pay off. But things are not so simple: Such government actions are likely to provoke retaliation. The U.S. government may retaliate by granting a \$10 billion subsidy to Boeing to produce the plane, with the result that both firms may decide to produce it, although this is not economically desirable. In fact, the European Union has charged that the United States subsidizes Boeing with defense contracts because much of the U.S. flight research and applications are transferable from military to commercial aircraft.

GOVERNMENT PRICE CEILINGS AND PRICE FLOORS

Government officials may intervene in markets domestically by enforcing prices that would not result had market forces been allowed to determine prices. We see this in the form of **price floors**, where the government will not allow a price to *fall* to its market level because of a belief or political pressure that the market-determined price is too low (minimum wage laws, agricultural price supports). We also see this in the form of **price ceilings**, where the government will not allow a price to *rise* to its market level because of a belief or political pressure that the market-determined price is too high (rent control).

We will again use our tools of producer and consumer surplus to evaluate the social welfare impacts of such government intervention in the market. Consider the impact of agricultural price supports. Figure 17.8 shows the market-clearing price (P_C) and quantity (Q_C) and the price floor (P_F) set by the government.

Suppose the government imposes the floor and suppliers, realizing that only Q_{DF} will be demanded, constrain themselves to produce only Q_{DF} . Consumer surplus is A and producer surplus (= variable profit) is total revenue ($B + G + I + J$) minus variable cost (J) or $B + G + I$. Social welfare under the price floor is $A + B + G + I$.

If the market price prevails, consumer surplus is $A + B + C$ and producer surplus is $G + H + I$, so social welfare is $A + B + C + G + H + I$. Social welfare is defined as consumer surplus + producer surplus + any government tax revenue − any government subsidy. The price floor decreases social welfare by $C + H$; this is called the deadweight loss triangle or social welfare triangle. **Deadweight loss** is defined as the difference between social welfare under perfect competition and social welfare under alternative pricing.

Price floors Where the government will not allow a price to fall to its market level because of a belief or political pressure that the market-determined price is too low.

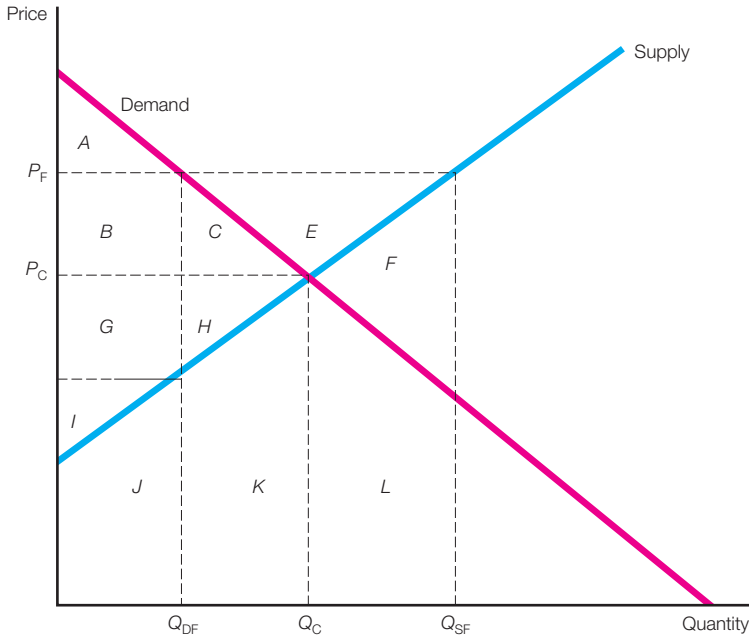
Price ceilings Where the government will not allow a price to rise to its market level because of a belief or political pressure that the market-determined price is too high.

Deadweight loss Social welfare under perfect competition minus social welfare under alternative pricing.

FIGURE 17.8

Impact of a Government Price Floor

A price floor decreases social welfare by the minimum of the deadweight loss triangle ($C + H$) or the maximum of the deadweight loss triangle plus the resource cost of producing the unused output ($F + K + L$).



In this case, consumers gain $A + B + C - A = B + C$ in consumer surplus by moving from a price floor to a market price. Producers gain $G + H + I - (B + G + I) = H - B$ by moving from a price floor to a market price. It's not clear whether producers gain or lose with a price floor in this case because $H - B$ can be positive or negative. Although it appears from Figure 17.8 that the producer will lose surplus if the price changes from the price floor to the market price, this is not necessarily true. A less elastic supply curve than the one drawn will increase H while leaving B unchanged and could make $H - B$ positive. Note the sum of the consumer and producer gain $B + C + H - B = C + H$ gives the social welfare gain in moving from a price floor to the market price. Thus the deadweight loss of a price floor in this case is $C + H$.

What if the suppliers produce what the supply curve tells them to do when the price is P_F , that is, produce Q_{SF} ? Because consumers will demand only Q_{DF} at that price, there will be a surplus of $Q_{SF} - Q_{DF}$ units in the market. Here's where the government comes in. The market buys Q_{DF} of the output at the price of P_F , and

the government purchases the rest at the market price for a government expenditure of $C + E + F + H + K + L$.

Consumer surplus with the price floor remains at A . Producer surplus is now all the producer revenue (including revenue from the government): $B + C + E + F + G + H + I + J + K + L$, minus the variable cost of producing Q_{SF} units ($F + J + K + L$), or $B + C + E + G + H + I$. Social welfare with the price floor is consumer surplus (A) plus producer surplus ($B + C + E + G + H + I$) minus government expenditure ($C + E + F + H + K + L$) or $A + B - F + G + I - K - L$. Social welfare with market pricing is $A + B + C + G + H + I$. Deadweight loss is $A + B + C + G + H + I - (A + B - F + G + I - K - L) = C + H + F + K + L$. Thus the deadweight loss with the government subsidy of buying the unsold production at the floor price increases by the cost of producing the unsold items. This is a wasteful use of scarce resources to produce unused output.

What if the government paid the producers not to produce the output that wouldn't be sold in the market? The government saves further expense by not having to transport, store, or destroy the output. How much should it pay? How about the producer surplus the producers will earn if they had been able to sell their $Q_{SF} - Q_{DF}$ at the price of P_F , that is, $C + E + H$? Consumer surplus with the price floor remains at A . Producer surplus is $B + G + I$ from sales to the market and $C + E + H$ from the government. Government expenditure is $C + E + H$. So social welfare with this version of the price floor is $A + B + G + I + C + E + H - (C + E + H) = A + B + G + I$. Social welfare with market pricing is $A + B + C + G + H + I$, and so the deadweight loss from this pricing floor scheme is $C + H$ —that is, the social welfare triangle. Thus if the government wants an agricultural support price floor, paying the farmers not to produce is the way to go.

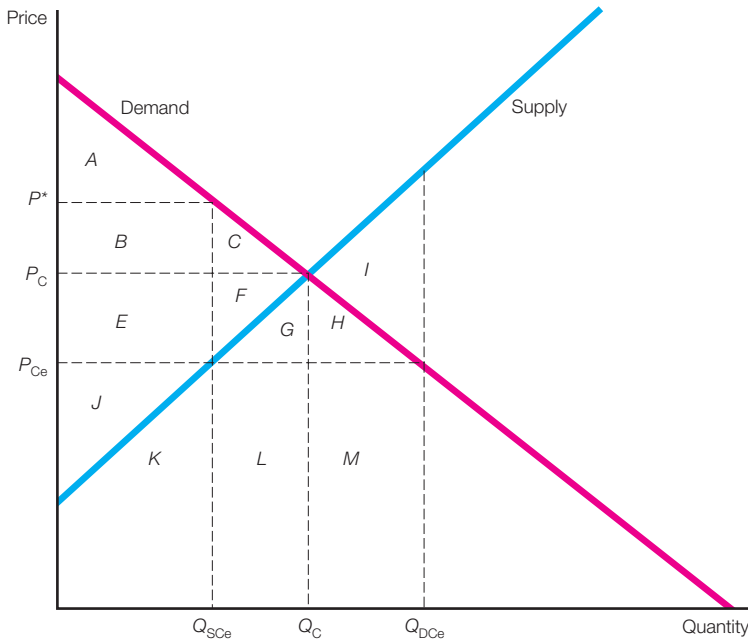
What about a price ceiling? Consider Figure 17.9. With a price ceiling of P_{Ce} , demanders will want $Q_{D_{Ce}}$ of the good, and suppliers will want to supply $Q_{S_{Ce}}$; so a shortage of $Q_{D_{Ce}} - Q_{S_{Ce}}$ exists. But because only $Q_{S_{Ce}}$ will appear on the market, that is the relevant figure. Those lucky enough to consume the good will receive consumer surplus of $A + B + E$. The producer will receive producer surplus of J . Social welfare will be $A + B + E + J$. If the market were allowed to function, the price would be P_C . Consumer surplus is $A + B + C$, producer surplus is $E + F + J$, and social welfare is $A + B + C + E + F + J$. The deadweight loss is $C + F$.

With a market price, consumers gain $A + B + C - (A + B + E) = C - E$ in consumer surplus. This can be either positive or negative and would appear to be negative in Figure 17.9. However, a much steeper demand curve through the market equilibrium coordinates will increase C while not changing E and potentially making $C - E > 0$. With a market price, producers gain $E + F + J - J = E + F$. Social welfare thus increases by $C - E + E + F = C + F$ with market pricing.

FIGURE 17.9

Impact of a Government Price Ceiling

A price ceiling decreases social welfare by the deadweight loss triangle $(C + F)$, if the government compensates producers for any loss they incur, or by the deadweight loss triangle I .



What is the significance of P^* in Figure 17.9? With Q_{DCe} demanders trying to get Q_{SCe} units, a rationing mechanism has to be used. This mechanism could be a lottery where the lucky Q_{SCe} winners get the units. It could also be a sneaky price mechanism. For example, it's been reported that people pay "key money" (equal to $P^* - P_{Ce}$) so that quantity demanded equals the quantity of units supplied at P_{Ce}) to obtain rent-controlled apartments.

Suppose the government wants to satisfy all the demand at price P_{Ce} . Producers will satisfy all the quantity demanded and receive the ceiling price for each unit. The government will then compensate the producers for any loss they incur. Consumer surplus is $A + B + C + E + F + G + H$. Producer surplus is $J + K + L + M + I + G + H - (I + G + H + K + L + M) = J$. The government expenditure is $I + G + H$ for the producer costs not covered by the revenues from consumer purchases. Social welfare is $A + B + C + E + F + G + H + J - I - G - H = A + B + C + E + F + J - I$. The deadweight loss triangle in this case is I , where the additional $Q_{DCe} - Q_C$ items consumed over

the optimal number bring benefit to consumers of $H + M$ but cost $I + H + M$ to produce.

An old TV ad stated, “It’s not nice to fool with Mother Nature.” Likewise, it’s not nice to fool with market outcomes (except when markets fail, such as with externalities, as we will point out soon).

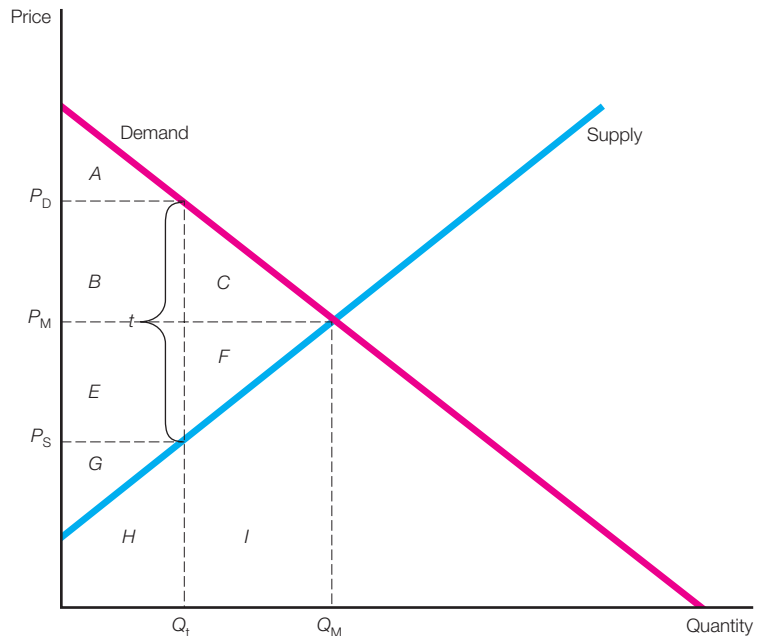
THE WELFARE IMPACTS OF TAXES

Suppose the government imposes a per-unit tax of t on a good and requires sellers to pay this tax to the government. This tax drives a wedge of magnitude t between the price a seller receives and what a demander pays. Consider Figure 17.10. If there is no tax, the market price prevails, and consumer surplus is $A + B + C$, producer surplus is $E + F + G$, and social welfare is $A + B + C + E + F + G$. With the tax, the demander pays a price of P_D and consumes Q_t . On net, the supplier receives $P_S = P_D - t$. Consumer surplus is A , producer surplus is G , govern-

FIGURE 17.10

The Incidence and Welfare Costs of a Per-Unit Tax

A per-unit tax of t causes a deadweight loss of $C + F$. The tax is borne more by the buyer or the seller depending on their relative own-price elasticities, with the least elastic bearing more of the tax.



ment tax revenues are $t \times Q_t = B + E$, and social welfare is $A + B + E + G$. The deadweight loss caused by the tax is $C + F$.

Who bears the brunt of the tax: the producer or the consumer? It depends on the relative elasticities of demand and supply. The elasticity of demand and supply through the market price equilibrium coordinates depends on the slope of the curves (recall from Chapter 2 that one definition of elasticity is $\eta = (P_M/Q_M)(1/\text{slope})$) because P_M and Q_M would be the same for all curves. A more gently sloped (more elastic) demand curve than the one depicted would decrease output and yield a lower P_D than the one depicted for a given t . Thus the demander would bear less of the tax with a more elastic demand curve. A steeper (less elastic) supply curve than the one depicted would increase output and yield a lower P_S than the one depicted for a given t . Thus the supplier would bear more of the tax with a less elastic supply curve. *The burden of the tax goes in the direction of the less elastic market participant.*

REGULATION OF ENVIRONMENTAL POLLUTION

Having looked briefly at antitrust, trade, tax policy, and the patent system, we return to the topic of government regulation. Government agencies regulate many aspects of economic life, not just the prices charged by public utility companies. Managers of firms in a wide variety of industries, ranging from steel or chemicals to paper or petroleum, must understand and cope with a huge number of government regulations to protect the environment. To illustrate the situation, consider the Reserve Mining Company (now Northshore Mining), which produces iron pellets from taconite rock. For every ton of iron pellets it manufactures, Reserve also produces two tons of waste taconite tailings, which for over a decade were dumped into Lake Superior. In 1969 Reserve found itself in a court battle, one of the most hotly debated matters being the discovery of asbestos-like fibers in the water supply of Duluth, Minnesota. When the legal battle was resolved in 1977, Reserve was granted the necessary permits to begin construction of a new dumping facility, which cost about \$400 million. The price was high, but asbestos levels in Lake Superior have dropped substantially.

In the following sections we explain why our economy, in the absence of government action, is likely to generate too much pollution. Then we discuss the optimal level of pollution control and describe various forms of government regulation.

External Economies and Diseconomies

To understand why our economy is likely to generate too much pollution, we must define an external economy and an external diseconomy. An **external economy** occurs when an action by a firm or individual gives uncompensated benefits to others. For example, a firm may train workers who eventually go to work for other

External economy When an action by a firm or individual gives uncompensated benefits to others.

firms, which need not pay the training costs. Or managers may carry out research that benefits other firms, which need not pay for the research. In general, there is a tendency for activities resulting in external economies to be underperformed from society's point of view. A firm or individual that performs an action that contributes to society's welfare but receives no payment for it is likely to perform this action less frequently than would be socially desirable.

External diseconomy When an action by a firm or individual results in uncompensated costs or harm to others.

An **external diseconomy** occurs when an action by a firm or individual results in uncompensated costs or harm to others. For example, a firm may generate smoke that harms neighboring families and businesses, or a person may fail to maintain his or her property, reducing the value of neighboring houses. In general, there is a tendency for activities resulting in external diseconomies to be overperformed from society's point of view. A firm or individual that performs an action that results in costs borne by others is likely to perform this action more frequently than is socially desirable.

The Genesis of the Pollution Problem

The key to understanding why our economy generates too much pollution (from society's point of view) is the concept of an external diseconomy. Firms and individuals that pollute our waterways and atmosphere are engaged in activities resulting in external diseconomies. They may pollute a river by pumping out waste materials, or pollute the air with smoke or other materials. These activities generate external diseconomies. Those that pollute (without penalty) transfer the costs of pollution to others; and as pointed out, they are likely to overpollute from a social viewpoint.

In a competitive economy, resources tend to be used in their socially most valuable way because they are allocated to the people and firms that find it worthwhile to bid most for them, assuming that prices reflect true social costs. Suppose, however, that because of the presence of external diseconomies, people and firms do not pay the true social costs for certain resources. In particular, suppose some firms or people can use water or air for nothing, but other firms or people incur costs as a consequence of this prior use. In this case, the price paid by the user of water or air is less than the true cost to society. In a case like this, users of water and air are guided in their decisions by the prices they pay. Because they pay less than the true social costs, water and air are artificially cheap for them, so that they use too much of these resources from society's point of view.

The Optimal Level of Pollution Control

Managers, like other members of society, should be able to look at matters from a social, as well as private, standpoint. They should be sensitive to the effects of their actions on society as a whole, as well as on their firm's interests. An industry generally can vary, at each level of output, the amount of pollution it generates. For

instance, it may install pollution control devices like scrubbers to cut the amount of pollution it generates. In this section we determine the socially optimal level of pollution control.

The total social cost of each level of discharge of an industry's wastes, holding constant the industry's output, is shown in Figure 17.11. The more untreated waste the industry discharges into the environment, the greater are the total costs. Figure 17.12 shows the costs of pollution control at each level of discharge of the industry's wastes. The more the industry reduces the amount of wastes it discharges, the higher are its costs of pollution control. Figure 17.13 shows the sum of these two costs (the cost of pollution and the cost of pollution control) at each level of discharge of the industry's wastes.

From society's point of view, the industry should lower its discharge of pollution to the point where the sum of these two costs (the cost of pollution and the cost of pollution control) is at a minimum. Specifically, the optimal level of pollution in the industry is *B* in Figure 17.13. To see why this is the optimal level, note that if the industry discharges less than this amount of pollution, a one-unit increase in pollution lowers the cost of pollution control by more than it increases the cost of pollution; whereas if the industry discharges more than this amount of pollution, a one-unit reduction in pollution lowers the cost of pollution by more than it increases the cost of pollution control.

FIGURE 17.11

Pollution Cost

The cost of pollution increases as larger quantities of pollutants are emitted.

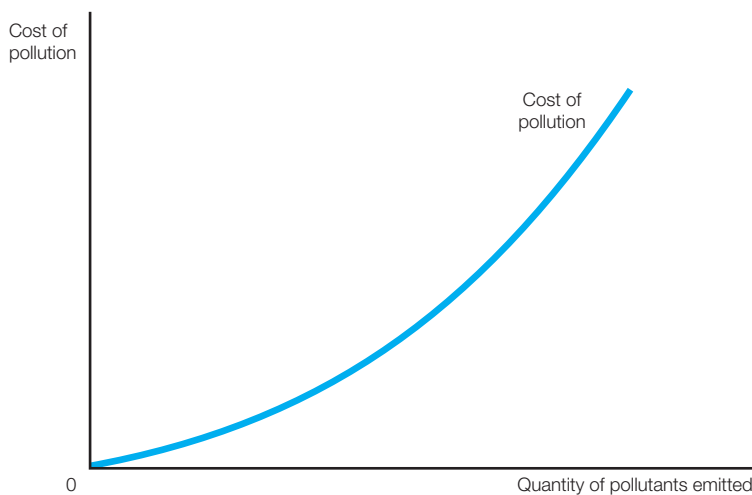


FIGURE 17.12

Pollution Control Cost

The cost of pollution control decreases as larger quantities of pollutants are emitted.

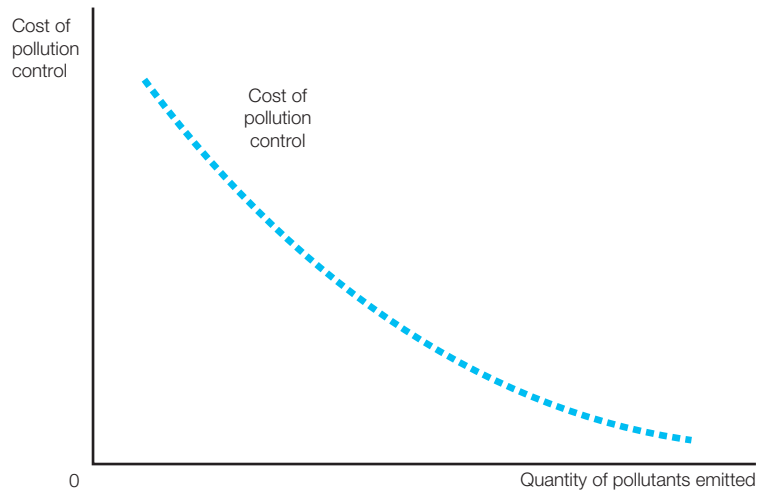


FIGURE 17.13

Sum of the Pollution Cost and the Pollution Control Cost

From the point of view of society as a whole, the optimal level of pollution in this industry is B .

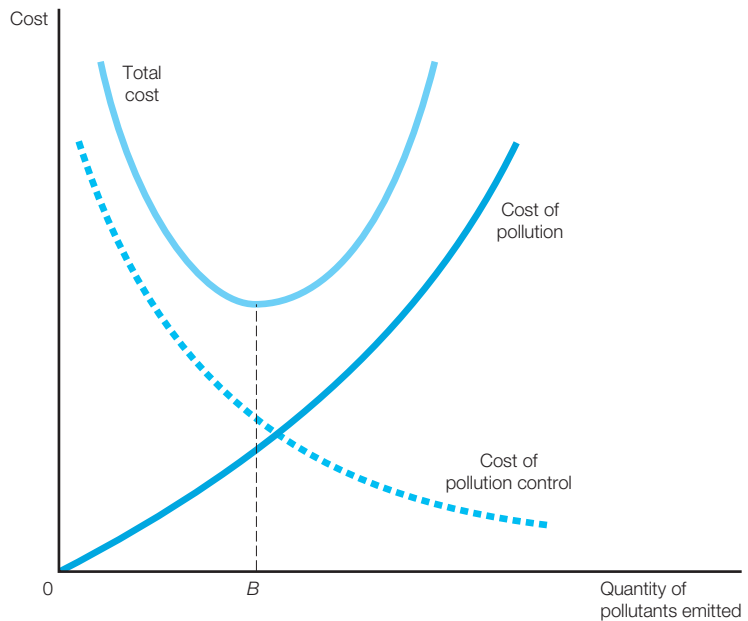


FIGURE 17.14

Marginal Cost of Pollution and Marginal Cost of Pollution Control

At the socially optimal level of pollution, B , the cost of an extra unit of pollution is equal to the cost of reducing pollution by an extra unit.

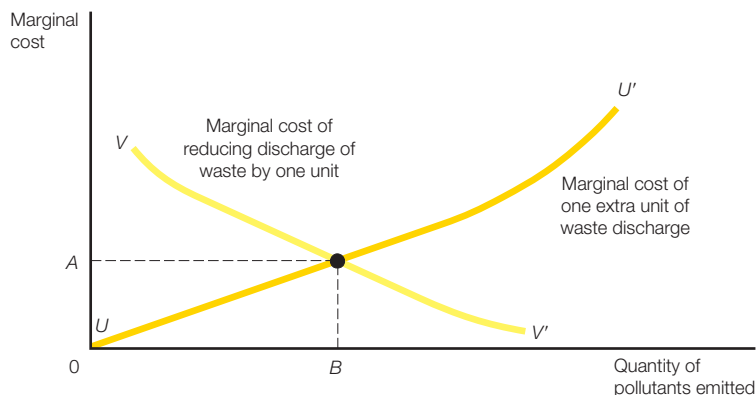


Figure 17.14 shows the marginal cost of an extra unit of discharge of waste at each level of discharge of the industry's wastes; this is designated UU' . Figure 17.14 also shows the marginal cost of reducing the industry's discharge of waste by one unit; this is designated by VV' . The socially optimal level of pollution for the industry is at the point where the two curves intersect. At this point, the cost of an extra unit of pollution is just equal to the cost of reducing pollution by an extra unit. Regardless of whether we look at Figure 17.13 or 17.14, the answer is the same: B is the socially optimal level of pollution.

Forms of Government Regulation

Because it does not pay all the social costs of its pollution, the industry in Figure 17.14 does not find it profitable to reduce its pollution level to B . One way the government can establish incentives for managers to reduce their pollution is by direct regulation. For example, the government may decree that this industry is to limit its pollution to B units. Direct regulation of this sort is relied on in many sectors of the U.S. economy.

Another way to induce managers to reduce pollution is to establish effluent fees. An **effluent fee** is the fee a polluter must pay to the government for discharging waste. For instance, in Figure 17.14 an effluent fee of A per unit of pollution discharge might be charged. If so, the marginal cost of an additional unit of pollution discharge to the industry is A , with the result that it cuts back its pollution to the socially optimal level, B units. To maximize their profits, the managers in the industry reduce pollution to B units because it is profitable to cut back pollution

Effluent fee The fee a polluter must pay to the government for discharging waste.

so long as the marginal cost of reducing pollution by a unit is less than A —and, as we see from Figure 17.14, this is the case when the pollution discharge exceeds B .

To illustrate the usefulness of effluent fees, consider Germany's Ruhr Valley, a highly industrialized area with limited water supplies. Effluent fees have been used in the Ruhr to help maintain the quality of the local rivers, and the results have been highly successful. But direct regulation is useful, too. Some ways of disposing of certain types of waste are so dangerous that the only sensible thing to do is to ban them. Also, it sometimes is not feasible to impose effluent fees—for example, when it is difficult to meter the amount of pollutants emitted by various firms and households.

Yet another way that the government can reduce the amount of pollution is to issue **transferable emissions permits**, which are permits to generate a particular amount of pollution. These permits, which are limited in total number so that the aggregate amount of pollution equals the level the government decides on, are first allocated among firms. They then can be bought and sold. Managers who find it expensive to curb pollution are likely to buy these permits; managers who find it cheap to do so are likely to sell them. The Clean Air Act of 1990 called for the use of such permits to reduce the emission of sulfur dioxide, and the Chicago Board of Trade voted to create a market for these permits (see problem 2 at the end of this chapter). An exchange for international carbon dioxide permits was formed in Amsterdam in 2005.

Transferable emissions permits
Permits to generate a particular amount of pollution.

Consider another way of solving the problem, attributable to Ronald Coase, who won the 1991 Nobel Prize for his contributions (including this one) to economics. The Haddonfield Brewery (a microbrewery) is located downstream from the Cherry Hill Chemical Company. Cherry Hill dumps its effluent into the river. Haddonfield requires clean water to brew its beer and must therefore filter the river water before using it. This filtering costs Haddonfield \$50,000, and its current profits are \$200,000. Cherry Hill's current profits are \$500,000. It is possible for Cherry Hill to refine the effluent before discharging it into the river so that the river water meets all standards for beer brewing. Installing and operating such a refining system would cost Cherry Hill \$40,000. Coase viewed the process of controlling pollution and its costs as one that could be solved without government intervention.

Suppose there are no government laws regarding pollution. Cherry Hill can pollute at will. The value of society's output as we initially view this situation is \$700,000, or $\$200,000 + \$500,000$.

Note that if Haddonfield managers paid \$40,000 to Cherry Hill to refine its effluent, Haddonfield could save \$10,000, increase its profits to \$210,000, and hence increase the societal welfare to \$710,000. The pollution is eliminated and the societal welfare is increased, and only private entities are involved.

However, such a solution could be thwarted by negotiation costs. Suppose negotiation costs for an agreement between Haddonfield and Cherry Hill are

\$11,000. Cherry Hill managers have no reason to negotiate or pay negotiation costs. Haddonfield managers would gain \$10,000 with an agreement, but they must bear \$11,000 in negotiation costs, so they will choose not to do so. It would be cheaper for them to filter the incoming water.

Suppose we impose the basic common law on this situation—not a law against pollution per se, but a law that says one is liable for the damage one causes to another. Cherry Hill's effluent causes \$50,000 in damages to Haddonfield (if Haddonfield had clean water, it would not have to spend \$50,000 on filtering and its profit would increase to \$250,000). Given liability for the damages it causes, Cherry Hill could shut down and stop polluting. But it is very profitable, and shutting down would be foolish because it would deprive society of its output. A better solution would be for Cherry Hill to pay Haddonfield \$50,000 for the damages caused; it would still have \$450,000. Still a better solution would be to install the refining system and incur a cost of \$40,000; Cherry Hill would still have \$460,000. This eliminates any liability on Cherry Hill's part because the water is no longer polluted.

Note that in the first case, where there was no law (and no negotiation costs) and Cherry Hill could do as it wished, the ultimate solution was for Haddonfield to pay Cherry Hill to put in water refining equipment. In the latter situation, where there was a liability law, the ultimate solution was for Cherry Hill to install the refining equipment with no payment from Haddonfield. With or without the law, the private sector chooses the cheapest method to eliminate the pollution problem, that is, Cherry Hill installing the refining equipment. In either case the societal welfare is \$710,000.

The difference between the two situations is the distribution of welfare between Haddonfield and Cherry Hill. With liability, Cherry Hill gets \$500,000 – \$40,000 = \$460,000 and Haddonfield gets \$250,000. With no liability, Cherry Hill gets \$500,000 + P – \$40,000 = \$460,000 + P and Haddonfield gets \$250,000 – P , where \$50,000 > P > \$40,000. Previously we assumed that Haddonfield paid Cherry Hill's cost of installing the refining equipment; but with no liability laws, Cherry Hill need not abate its pollution. Clearly Cherry Hill managers must receive at least \$40,000 (the cost of the refining equipment) from Haddonfield; because getting rid of the pollution is worth \$50,000 to Haddonfield, Cherry Hill managers could hold out for a payment P greater than \$40,000 (but Haddonfield would never pay more than \$50,000 because it can do its own filtering for \$50,000).

Therefore, under the no-liability scenario, Cherry Hill would get between \$500,000 and \$510,000, and Haddonfield would get between \$210,000 and \$200,000. Cherry Hill prefers the nonliable scenario and Haddonfield prefers liability; but either way, society gets social welfare of \$710,000, and pollution is abated in the cheapest way.

STRATEGY SESSION: Buying and Selling the Right to Emit Greenhouse Gases

The first trading of emissions permits for an air pollutant that affects human health occurred in 1994 when Public Service Enterprise Group (NJ) reduced its emissions of nitrogen oxides by 2,400 more tons than required by law, and Northeast Utilities (a New England utility operator) announced that it intended to purchase the right to emit 500 of those tons. Since that time, we've come a long way.

The Chicago Climate Exchange (CCX) was North America's largest and longest running greenhouse gas emission reduction program. It started operations in 2003 as a cap-and-trade program. It ended operations as of December 31, 2010. Operations in Europe (European Climate Exchange) continue.

A synopsis of trade volumes (rights to pollute in metric tons) and prices (per metric tons) paid for a permit to emit is discussed below. There was a spike in volume and price before the economy tanked in late 2008 and the very low volumes and prices in the depth of the recession and early recovery resulted. At the peak, rights to pollute a metric ton were about \$7.50. By July 2011, there were only four transactions: one for 500 metric tons for \$0.75 per metric ton

on July 11 and three on July 1 for 1,500 metric tons at \$0.60 per metric ton as futures transactions were fulfilled. While most transactions involved the United States, in transactions dating back to July 2010, two involved India, with one each involving Brazil, Uruguay, and China.^a

The emission permit markets in Europe are also showing the impacts of the slow economy. *Bloomberg News* reports that emission allowances for the European Union for December 2010 fell 2% to \$24.08 per metric ton after the EU Central Bank president talked about slow economic growth.^b

The New Zealand Emissions Trading Scheme (enacted into law in 2008) entails the trading of New Zealand Units (NZUs) among entities. The purpose of the trading scheme is to reduce the amount of greenhouse gases emitted in the country. Those who emit such gases, such as mines, will be charged when they emit the gases. They will have to "pay" NZUs to the government in order to emit greenhouse gases. Others, such as forests, will earn NZUs from the government because their activities absorb greenhouse gases. In some cases, the government will give NZUs

Effects of the Regulation-Induced Cost Increase on Price and Output

Regardless of how the government induces firms to reduce pollution, the result is an increase in the firms' costs, as in the case of the Reserve Mining Company discussed earlier. Firms spend substantial amounts annually on environmental costs (both in prevention and in penalties for exceeding legal limits); for example, DuPont is reported to have spent about \$572 million (pretax) on environmental expenses in 2009.⁹ It is important to recognize this and to see how to determine the extent to which this cost increase is passed on to consumers in the form of a price increase—and the extent to which it will be borne by the firms. In this section we learn how this can be done. It is analogous to the incidence of a tax shown earlier.

Suppose a new regulation says that paper mills must use new methods to reduce water pollution. Assuming the paper industry is perfectly competitive, we

9. DuPont, 2010 Global Reporting Initiative Report, p. 43, at www2.dupont.com/Sustainability/en_US/assets/downloads/gri.pdf.

to an entity if the government determines that a greater social good will occur by such a distribution. During the initial phase of implementation (July 1, 2010 to December 31, 2012), participants can purchase NZUs from the government for \$25, and those in the energy, industrial, and liquid fossil fuel sectors will surrender one NZU for each two metric tonnes of emissions produced. In 2013 and beyond, the surrender rate will be one NZU for each tonne emitted and the price of an NZU will be market determined (as opposed to a set \$25).

The scheme puts a price on greenhouse gas emissions, and this price creates an incentive for entities that emit to reduce those emissions (the impact of a downward sloping demand curve). At the same time, it creates an incentive to plant forests to absorb carbon dioxide. With emissions trading, participants can choose how they wish to comply with their obligations to reduce emissions; that is, they can reduce production of their primary product, change their production function to one that is more emission friendly, or purchase NZUs that enable them to emit more.

The NZUs are for carbon dioxide (CO_2), methane (CH_4), nitrous oxide (N_2O), hydrofluorocarbons (HFCs), perfluorocarbons (PFCs), and sulphur hexafluoride (SF_6).

The scheme will be administered by the Ministry of Economic Development, which will run the day-to-day operations. It is also responsible for insuring that participants are complying with the scheme. New Zealand views the trading scheme as the government's primary response to global climate change. Note also that New Zealand is a signatory to the Kyoto Protocol.^c

^aSee www.theice.com/ccx.jhtml and see Nathaniel Gronewold, "Chicago Climate Exchange Closes Nation's First Cap-and-Trade System but Keeps Eye to the Future," *New York Times*, January 3, 2011.

^bSee Catherine Arlie, "European Carbon Permits Fall as Slow Economy May Sap Demand," *Bloomberg.com*, May 6, 2011, at www.bloomberg.com/news/print/2011-05-06/european-carbon-permits-fall-as-slow-economy-may-sap-demand.html.

^cSee "Emissions Trading Scheme basics" at www.climatechange.govt.nz/emissions-trading-scheme/about/basics.html.

can compare the situation in the industry after the regulation's enactment with that prior to its enactment. Before the regulation, the marginal cost function of each paper producer is assumed to be

$$MC = 20 + 40Q \quad (17.11)$$

where Q equals the number (in thousands) of tons of paper produced per week. If the price is P , managers, to maximize profit, set price equal to marginal cost, which means that

$$P = 20 + 40Q$$

or

$$Q = -0.5 + 0.025P$$

If there are 1,000 paper producers, all with the same cost function, the industry's supply curve is

$$Q_S = 1,000(-0.5 + 0.025P) = -500 + 25P \quad (17.12)$$

Assuming that the market demand curve for paper is

$$Q_D = 3,500 - 15P \quad (17.13)$$

we can find the equilibrium price and output of paper by setting the quantity demanded in equation (17.13) equal to the quantity supplied in equation (17.12)

$$\begin{aligned} 3,500 - 15P &= -500 + 25P \\ 40P &= 4,000 \\ P &= 100 \end{aligned}$$

The quantity demanded equals

$$Q_D = 3,500 - 15P = 3,500 - 15(100) = 2,000$$

And the quantity supplied equals the same amount

$$Q_S = -500 + 25P = -500 + 25(100) = 2,000$$

In other words, before the new regulation, the price of paper is \$100 per ton, and 2,000 tons are produced per week.

What is the effect of the new regulation on the price and output of paper? Suppose the regulation raises the marginal cost of producing paper by 25%. After the regulation, the marginal cost function of each paper producer is

$$\begin{aligned} MC &= 1.25(20 + 40Q) \\ &= 25 + 50Q \end{aligned}$$

To maximize profit, each firm sets marginal cost equal to price, which means that

$$25 + 50Q = P$$

or

$$Q = -0.5 + 0.02P$$

Hence the industry's postregulation supply curve is

$$\begin{aligned} Q'_S &= 1,000(-0.5 + 0.02P) \\ &= -500 + 20P \end{aligned} \quad (17.14)$$

if all 1,000 paper producers stay in the industry. (Some may drop out if they cannot avoid losses—recall Chapter 7.) To find the equilibrium price after the enactment of the new regulation, we set the quantity demanded in equation (17.13) equal to the quantity supplied in equation (17.14)

$$\begin{aligned} 3,500 - 15P &= -500 + 20P \\ 35P &= 4,000 \\ P &= 114.29 \end{aligned}$$

PROBLEM SOLVED: Auctioning Off Spectrum Rights: Australia and the United States

In 1995 the Federal Communications Commission (FCC) completed the first auction for broadband personal communications services (PCS) licenses. In designing its auction, both the FCC and its economic advisers were aware that auctions could backfire, as illustrated by a famous case in Australia. In April 1993 two licenses for satellite television service were auctioned off in Australia. When the sealed bids were received, the winners were Hi Vision Ltd. and Ucom Pty. Ltd.; their winning bids were about \$140 million and \$120 million, respectively. Because these bids were larger than expected and because these firms were not among the major players in the Australian television industry, the Australian government announced that the auction ushered in a “whole new era.”

To the government’s dismay, both Hi Vision and Ucom defaulted on their highest bids. Hence the licenses had to be reawarded at the next highest levels, which were also theirs. It soon became clear that each firm had submitted many bids, each about \$5 million higher than the next. After defaulting on a number of its bids, Ucom eventually paid about \$80 million for one license and \$50 million for the other. An Australian politician called it “one of the world’s great media license fiascos,” and Bob Collins, Australia’s communications minister, was almost fired.

Auctions can also be tremendously successful. In March 2008 the U.S. government announced that companies bid more than \$19 billion for the portion of the wireless spectrum that was freed as the result of the conversion to digital television by broadcasters. This was the most lucrative auction in history.

Problems

1. What are the advantages of auctions over other schemes to choose who gets licenses?
2. What was the fundamental flaw in Australia’s auction of licenses?
3. To help avoid such flaws, the FCC stipulated that firms had to make down payments to the FCC; and if a high bidder were to withdraw its bid during the auction, it would be liable for the difference between its bid and the price ultimately obtained for the license. (For bids withdrawn after the auction, there would be a supplementary penalty of 3%.) Why would these stipulations help avoid such flaws?

Solutions

1. Auctions tend to reduce the costs and delays in choosing licenses. Hearings and lotteries, which are the principal alternative ways to allocate licenses, use a great deal of resources, particularly the time of economic, engineering, and legal consultants.
2. The fundamental flaw in Australia’s auction was the lack of a penalty for default, which implied that bids were not meaningful.
3. These stipulations were aimed at preventing bidders from defaulting without penalty on their bids.^a

^aFor further discussion, see John McMillan, “Selling Spectrum Rights,” *Journal of Economic Perspectives*, Summer 1994; and Stephen Labaton, “Wireless Spectrum Auction Raises \$19 Billion,” *New York Times*, March 19, 2008, at www.nytimes.com/2008/03/19/technology/19fcc.html.

Hence the postregulation quantity demanded equals

$$Q_D = 3,500 - 15P = 3,500 - 15(114.29) = 1,785.71$$

And the quantity supplied equals the same amount

$$Q'_S = -500 + 20P = -500 + 20(114.29) = 1,785.71$$

In other words, after the new regulation, the price of paper is \$114.29 per ton, and 1,785.71 thousands of tons of paper are produced per week.

Clearly the new regulation increases price (from \$100 to \$114.29 per ton—but note by less than the 25% cost increase) and reduces output (from 2,000 to 1,785.71 thousands of tons per week). This typically is the effect of such regulations, but the extent of the price increase (and the output reduction) depends on the price elasticity of demand for the product. If the price elasticity is low in absolute value, the price increase is greater (and the output reduction smaller) than if the price elasticity is high in absolute value.

PUBLIC GOODS

In addition to regulating the environment and the behavior of monopolists, the government performs various economic functions, including providing goods and services. For example, the government is responsible for national defense, a critically important product in any society. Why does the government provide some goods and not others? One important reason is that some goods—so-called public goods—are unlikely to be produced in sufficient amounts by the private (nongovernmental) sector of the economy. Therefore the government is given the task of providing these goods. Before concluding this chapter, we must describe briefly what a public good is and why the private sector is unlikely to provide a public good in sufficient amounts.

A major hallmark of a public good is that it can be consumed by one person without diminishing the amount that other people consume of it. Public goods tend to be relatively indivisible; they often come in such large units that they cannot be broken into pieces that can be bought or sold in ordinary markets. Also, once such goods are produced, there is no way to bar citizens from consuming them. Whether or not citizens contribute toward their cost, they benefit from them. Obviously this means that it would be difficult for any firm to market them effectively.

National defense is a public good. The benefits of expenditure on national defense apply to the entire nation. Extending the benefits of national defense to an additional citizen does not mean that any other citizen gets fewer of these benefits. Also, there is no way of preventing citizens from benefiting from them, whether they contribute to the cost or not. Therefore ordinary markets (such as those for wheat, steel, or computers) cannot provide national defense. Because it is a public good, national defense, if it is to reach an adequate level, must be provided by the government; the same is true for flood control, environmental protection, and other such services.

Although these services are provided by the government, this does not mean they must be produced entirely by the government. The U.S. Air Force does not produce the B-2 Stealth Bomber; Northrop Corporation (now part of Northrop Grumman) does. The U.S. Navy does not produce the F-14 fighter; Grumman

STRATEGY SESSION: Entrance Fees to National Parks

The United States has many national parks, such as Yellowstone, and people frequently complain that they are overcrowded. Some economists, such as Allen Sanderson of the University of Chicago, have suggested that the solution is to raise entrance fees. They point out that when the National Park Service was set up in 1916, a family of five arriving by car could gain admission to Yellowstone for \$7.50; in 1995 the price was only \$10. If the 1916 price had risen in accord with the general rate of inflation, the 1995 fee would have been about \$120.

According to Sanderson, “We are treating our national and historical treasures as free goods when they are not. We are ignoring the costs of maintaining these places and rationing by congestion—when it gets too crowded, no more visitors are allowed—perhaps the most inefficient way to allocate scarce resources.”^a

The U.S. National Parks service has responded by charging admission to its parks. Here we show a schedule of fees taken from each park’s website in the summer of 2011. Although the prices don’t approach Sanderson’s inflation-adjusted Yellowstone fee, they now serve as a rationing device for limited park capacity.

The U.S. National Parks Service offers a pass (America the Beautiful—National Parks and Federal Recreational Lands Pass), in cooperation with the U.S. Department of Agriculture—Forest Service, Fish and Wildlife Service, Bureau of Land Management, and Bureau of Reclamation, that allows a vehicle, the pass holder and three other adults, and an unlimited number of children under the age of 16 access to facilities run by the agencies for a year for \$80. A senior citizens’ version costs \$10, and admission is free for citizens with permanent disabilities and park volunteers.

Yearly memberships are also available. In 2011 there will be 17 free days, for example, Martin Luther King Jr. weekend.

Here are some general entrance fees for national parks for a noncommercial vehicle and its occupants:

Acadia	Seven-day permit for auto and occupants: \$20
Arches	Seven-day permit for auto and occupants: \$10
Bryce Canyon	Seven-day permit for auto and occupants: \$25
Death Valley	Seven-day permit for auto and occupants: \$20
Grand Canyon	Seven-day permit for auto and occupants: \$25
Grand Teton	Seven-day permit for auto and occupants: \$25 Seven-day permit for walk-ins, bike-ins: \$10 Seven-day permit for motorcycles: \$15 Year permit: \$50 (All allow entry to Yellowstone too.)
Great Smoky Mountains	Free
Yellowstone	Seven-day permit for auto and occupants: \$25
Yosemite	Seven-day permit for auto and occupants: \$20
Zion	Seven-day permit for auto and occupants: \$25

^a Allen Sanderson, “Save the Parks and Make a Profit” Op-Ed in *New York Times*, September 30, 1995, p. 19.

STRATEGY SESSION: All That Glitters Is Not Gold

In Chapter 1, we saw how commodity prices increased in 2008. By 2009, the world found itself mired in the biggest recession since 1929. This caused commodity prices to plunge. From an all-time high of \$147/bbl in July of 2008, oil hit a four-and-a-half year low at \$34/bbl in 2009. Consumers stopped spending; lay-offs increased on a global scale. The credit markets froze. The cold state was not from lacking liquidity as governments across the globe stimulated their economies. What was lacking was trust. That fragile gift managers sometimes bestow upon each other in the name of business. It seems the only institution trusted in 2009 was the U.S. government. People were actually paying the government to hold their money for them. Managers should never forget this lesson regarding the benefits of trust. It is necessary for those who value efficiency.

The day of free market reckoning has come and past. For the first time in our lives, we better understand what it really means to say the U.S. government is the lender of last resort. It seems it is the only lender in some business sectors. In great market movements, there are always unintended consequences. We offer this. Prior to the crisis, if we went to cocktail parties and talked about adverse selection and moral hazard, people would change the subject or get up and leave. Now, this is what many want to discuss. We could not be more delighted to better explain these concepts to them.

The crisis provides ample examples to illustrate the usefulness of this book to the managerial world. We measure usefulness by its ability to predict behavior. Let's look at moral hazard and adverse selection, the subjects of Chapter 15 and 16. One early response from the government to the financial crisis was called the Troubled Assets Relief Program (TARP). In its initial form the money from TARP was to be used to purchase illiquid assets (toxic) held on the books by financial institutions. Many of these assets

were mortgage-backed securities (MBS). These instruments, which were designed by investment bankers, pooled mortgages of different structure (rating, length). The pool would then be segregated into what are called tranches; tranches could have different ratings in terms of risk based on its position in the repayment schedule. In turn, these securities were included in pools with those of other credit markets, like credit cards or commercial loans, as collateral for another managerially-designed instrument called collateralized debt obligations (CDOs).

On October 3, 2008, Congress, by an overwhelming vote, gave \$700 billion to the Treasury Department to purchase illiquid (toxic) assets like MBSs and CDOs from financial institutions. One plan called for using reverse auctions to purchase the toxic assets and create a market price for them. Less than five weeks later on November 18, 2008, the Secretary of Treasury, Hank Paulson, announced the abandonment of the plan. Instead of purchasing toxic assets, the Treasury would now use fund monies to purchase equity stakes in financial institutions. What caused the Treasury to reverse its course of action? And, why were the assets toxic? These questions and many more are explained by looking at the impact of moral hazard and adverse selection on managerial behavior.

The synthetic products created by Wall Street bankers, and now illiquid, were poorly designed to protect investors from moral hazard induced behavior. This, in turn, created markets characterized by adverse selection. The securitization of mortgages (and other debt) brought imbalance to the market. Before securitization, most financial institutions held the mortgages they wrote. There was an incentive to correctly identify the appropriate level of risk because your institution would bear it. Now that chain of reasoning was broken. Mortgages written by the institution were packaged into bundles and sold to others. Managers of the firm that originated the loan could relax their

lending standards because they would quickly sell the loan to others. The bankers had created a situation conducive to moral hazard-like behavior.

The Impact of Moral Hazard-like Behavior

Situations predicted by moral hazard models became reality. Some managers compounded the impact of these principles through their actions. Managerial choices at Washington Mutual Savings (WaMu) typified such behavior.

Like most mortgage originators, those originated by managers at WaMu were quickly packaged and sold to others. But, managerial behavior created additional incentives for moral hazard-like behavior. For example, upper level managers were compensated on growth numbers. And, their bonus scheme did not consider bad debt due to nonperforming mortgages. Also, mortgage brokers were given greater commissions for writing more risky loans because such loans paid higher upfront fees. WaMu managers also pressured real estate appraisers to inflate property values in their reports. A founder of one such company said, "It was the Wild West. If you were alive, they would give you a loan. Actually, I think if you were dead, they would still give you a loan."^a WaMu became one of the biggest writers of sub-prime mortgages. Managers decided to focus on a variable rate product called option ARMs. These allowed the lender to choose how much to pay each month, and were offered with low initial interest rates. By 2006 these riskier mortgages comprised 70% of all mortgages written. Managers focused on these mortgages because they paid higher upfront fees, and allowed WaMu to book profits on the interest due, even if the borrower chose to defer paying it. In 2005 and 2006, Kerry Killinger, the CEO, received compensation totaling \$43 million due to these policies. Two years later WaMu suffered losses of billions of dollars due to bad loans and the company was sold at a great discount to JP MorganChase.

And, what about the bankers who were selling these bundled securities to investors? They also paid little consideration to reducing moral hazard-like behavior. And, their collective decisions helped to intensify the effect. They structured CDOs as private placement vehicles. The bonds were not subject to most regulatory bodies. Pricing was not transparent. There was never any public market for these products, and only those who traded them were privy to sale information. Investment banks got paid after the bonds were sold. Underwriting fees were estimated to be roughly 1.1%. Hence, for every \$1 billion in bonds sold, bankers would receive \$11 million in fees.^b Many bankers who sold the product did not even know what assets served as the underlying collateral. One said their stated goal was to "sell as many as possible and get paid the most for every bond sold."^c

Because the make-up of a CDO is opaque, investors depended on bond rating agencies to look more closely at the bond and rate it for risk. But, the bond raters were compensated for their services by the investment banks, and often gave AAA ratings to bonds containing sub-prime mortgages. Here, again, we see how managerial actions encourage moral hazard-like behavior.

Some bankers even described the moral hazards they were facing in documents to investors. Jill Drew describes one, "The Dillion Read fund also purchased \$45 million of preferred stock in the Mantoloking trust. The hedge fund's dual role as servicer and investor gave it the incentive to load the CDO with risky investments to enhance its potential return, according to the offering document."^d

The Impact of Adverse Selection

By late 2008, the market for MBSs and CDOs did not exist. The market had failed. Pricing the instruments was difficult because few managers understood the

(continued)

STRATEGY SESSION: All That Glitters Is Not Gold (Continued)

value of the underlying collateral. The government felt it had to intervene so it created TARP. The original idea for TARP was to help create price transparency. So, the Treasury initially thought about holding an auction to establish a price for these assets. The proposed design was based on reverse auctions. In these auctions there are many sellers (financial institutions), and one buyer (the Treasury). The sellers bid the price down. For these auctions the price reflects how many cents on the dollar managers were willing to sell their debt to the government. Managers who are willing to sell at the lowest cents on the dollar would get to sell debt to the government. The auctions would not only get bad debt off the books of some firms, it would also establish some basis for pricing these bonds.

The plan had some flaws, none more important than those caused by adverse selection. The managers who held the debt clearly knew more about its composition than the government, especially because benchmarks for relative comparisons of bonds were difficult. Few trusted the ratings bestowed by rating agencies given the subsequent loss in bond value and potential for moral hazard-like behavior. And, most bonds had a unique value given the variety of mortgages in a pool.

Treasury officials anticipated creating a “market for lemons.” Managers who owned the debt were like used car sellers, and the government played the used car buyer. The incentives were for the auction winners to saddle the government with their expected worst

(now also part of Northrop Grumman) does. Firms play a central role in developing and producing the weapons systems on which our military establishment relies, even though national defense is a public good.

Because the distinction between public and private goods is not always clear, Figure 17.15 gives a way of classifying goods.

Pure public goods have the characteristics of being nonrival (my consumption doesn’t impact your ability to consume) and nonexcludable (there is no mechanism to exclude people from access). A typical entry in the northwest cell of public good would be the previously mentioned national defense. Historically lighthouses have also been put in such a category. However, technology changes things; today all the information conveyed by a lighthouse can be transmitted electronically and with a scrambled signal, so excludability is now possible—the information has moved lighthouses from the northwest corner of Figure 17.15 to the southwest corner. Some goods, such as education, bridge the two cells because they have some, but not all, of the characteristics of pure public goods. As a result, public education is in the northwest cell and private education is in the southwest cell, and consumers choose whether they want to use “free” public education or pay market prices for private education.

Certain goods are nonexcludable but rival (the northeast cell). A public sidewalk is available to anyone, but it is subject to capacity constraints, and my occupying a spot may preclude you from being on the sidewalk. We used to think of

performing bonds. Because of adverse selection, the Treasury was destined to purchase “lemons.” The auctions might establish a price, but it would be a price for low-quality bonds.

This is one reason why Treasury officials changed their minds about how they would use TARP. The return on the monies was expected to be too low. Treasury officials recognized the dangers of adverse selection and chose to buy equity stakes in institutions themselves in return for some managerial control. The markets needed to stabilize and this was probably quicker and more transparent in showing government commitment.

In late 2008, Treasury officials did announce a program to purchase \$500 billion of MBSs by June

2009. In designing the program, officials did recognize the adverse selection issues and tried to mitigate them. Only securities guaranteed by Fannie Mae, Freddie Mac, or Ginnie Mae were considered for purchase. And, securities had to be fixed rate and could not include any interest-only bonds.

^aPeter S. Goodman, and Gretchen Morgenson; “Saying Yes, WaMu Built Empire on Shaky Loans,” *New York Times*, December 28, 2008, at www.nytimes.com/2008/12/28/business/28wamu.html.

^bDrew, Jill, “Frenzy,” *The Washington Post*, December 16, 2008, p. A01.

^cIbid, p. 3.

^dIbid, p. 6.

FIGURE 17.15

Classification of Goods

	Nonrival	Rival
Non-Excludable	Public goods	Common property
Excludable	Marketable public goods	Private goods

public roads as being analogous to sidewalks. Now tolls are being proposed for some existing free roads: High occupancy vehicle (HOV) lanes are being converted to HOT lanes (tolls charged for cars not meeting the vehicle occupancy criterion). Thus we have created ways to exclude consumers from some rival public goods. Finally, in the southeast corner of the matrix are private sector goods and services. Whether government provides goods in any or all of these cells is a decision made by the government or the people that elect the government. A trend toward privatization has led a number of governments to exit functions performed in the southeast corner, such as trash collection, airline operation, and prison operation.

Government agencies have a vital influence over a wide variety of industries, not just defense contractors like Northrop Grumman. This chapter has described in detail many activities of government agencies that are important to managers.

SUMMARY

1. Commissions regulating public utilities often set price equal to average total cost, including a fair rate of return on the firm's investment. One difficulty with this arrangement is that because the firm is guaranteed this rate of return (regardless of how well or poorly it performs), there is no incentive for managers to increase efficiency. Although regulatory lag creates some incentives of this sort, they often are relatively weak.

2. The Sherman Act outlaws any contract, combination, or conspiracy in restraint of trade and makes it illegal to monopolize or attempt to monopolize. The Clayton Act outlaws unjustified price discrimination and tying contracts that reduce competition, among other things. The Robinson-Patman Act was aimed at preventing price discrimination in favor of chain stores that buy goods in large quantities. The Federal Trade Commission Act was designed to prevent undesirable and unfair competitive practices.

3. The real impact of antitrust laws depends on the interpretation of these laws by the courts. In its early cases, the Supreme Court put forth and used the famous rule of reason—that only unreasonable combinations in restraint of trade, not all trusts, required conviction under the Sherman Act. The situation changed greatly in the 1940s, when the court decided that Alcoa, because it controlled practically all the nation's aluminum output, was in violation of the antitrust laws. In the early 1980s two major antitrust cases were decided against American Telephone and Telegraph and the IBM Corporation. The European Union has become very active in antitrust enforcement in the 2000s.

4. Patent laws grant an inventor exclusive control over the use of an invention in exchange for his or her making the invention public knowledge. The patent system enables inventors to obtain a larger portion of the social benefits from their inventions, but it frequently has only a limited effect on the rate at which imitators appear. Nonetheless, firms continue to make extensive use of the patent system.

5. Trade is a significant and growing portion of our economy. Government policy influences how free our trade is. This chapter discussed the perfect competition argument for free trade and showed the welfare impacts of tariff and quotas relative to free trade. In general, a tariff is a less harmful way to restrict trade than a quota.

6. Governments intervene with market-determined prices by imposing price floors (not allowing the price to fall to its equilibrium level) or price ceilings (not allowing the price to rise to its equilibrium level). In both situations the deadweight loss to society of such interference can be shown. A government tax on individual

goods drives a wedge between the price demanders pay and the price sellers retain. The welfare impacts of price interventions and taxes can be demonstrated using the same tools (producer and consumer surplus) developed in earlier chapters.

7. An external economy occurs when an action of a firm or individual gives uncompensated benefits to others. An external diseconomy occurs when an action of a firm or individual causes uncompensated costs or harm to others. Firms and individuals that pollute our waterways and atmosphere are engaged in activities resulting in external diseconomies.

8. The socially optimal level of pollution (holding output constant) is at the point where the marginal cost of pollution equals the marginal cost of pollution control. In general, this is at a point where a nonzero amount of pollution occurs. To formulate incentives that lead to a more nearly optimal level of pollution, the government can establish effluent fees, issue transferable emissions permits, or enact direct regulations, among other things. These policies internalize the externality and make production of the externality an implicit cost to the producer.

9. Regulations (and other measures) designed to reduce pollution tend to increase the costs of the regulated firms. The price of their products generally rises, and industry output tends to drop. If the price elasticity of demand is relatively low in absolute value, more of the cost increase can be passed along to consumers in the form of a price increase than would be the case if the price elasticity of demand were relatively high in absolute value.

10. A public good can be consumed by one person without diminishing the amount of it that other people consume. Also, once a public good is produced, there is no way to bar citizens from consuming it. Public goods, such as national defense, are unlikely to be produced in sufficient quantities by the private (non-governmental) sector of the economy. Therefore the government often is given the task of providing these goods.

PROBLEMS

1. In 1985 United Airlines purchased Pan Am's Pacific Division for \$750 million. The Department of Justice opposed the purchase, but it was approved by the U.S. Department of Transportation. The percentages of total passengers carried across the Pacific by each airline in 1984 were as follows:

Firm	Percentage	Firm	Percentage
Northwest	27.5	United	7.3
JAL	21.9	China Airlines	6.8
Pan Am	18.5	Singapore Airlines	2.9
Korean Air	9.3	Other	5.8

- a. What was the concentration ratio before the purchase? Was it relatively high?
 - b. What was the concentration ratio after the purchase?
2. The Chicago Board of Trade voted to create a private market for rights to emit sulfur dioxide. The Clean Air Act of 1990 established a limit, beginning in 1995, on total emissions of sulfur dioxide from 110 power plants. Firms finding it relatively expensive to cut their sulfur dioxide emissions are likely to buy pollution permits because such permits cost less than cutting their emissions. Given that firms can exceed their legal limits and pay fines of \$2,000 per ton, do you think that the price of a right to emit a ton of sulfur dioxide exceeds \$2,000? Why or why not?
 3. The Miller-Lyons Electric Company is engaged in a rate case with the local regulatory commission. The demand curve for the firm's product is

$$P = 1,000 - 2Q$$

where P is price per unit of output (in dollars) and Q is the output (in thousands of units per year). The total cost (excluding the opportunity cost of the capital invested in the firm by its owners) is

$$TC = 50 + 0.25Q$$

where TC is expressed in millions of dollars and Q is the output (in units per year).

- a. The Miller-Lyons Electric Company has requested an annual rate (that is, price) of \$480. If the firm has assets of \$100 million, what would be its rate of return on its assets if this request is granted?
 - b. How much greater would the firm's accounting profit be if it were deregulated?
4. The cost of pollution (in billions of dollars) originating in the paper industry is

$$C_P = 2P + P^2$$

where P is the quantity of pollutants emitted (in thousands of tons). The cost of pollution control (in billions of dollars) for this industry is

$$C_C = 5 - 3P$$

- a. What is the optimal level of pollution?
 - b. At this level of pollution, what is the marginal cost of pollution?
 - c. At this level of pollution, what is the marginal cost of pollution control?
5. Seven firms produce kitchen tables. Suppose their sales in the year 2011 are as follows:

Firm	Sales (Millions of Dollars)
A	100
B	50
C	40
D	30
E	20
F	5
G	5

- a. What is the concentration ratio in this industry?
 - b. Would you regard this industry as oligopolistic? Why or why not?
 - c. Suppose that firm A merges with firm G. What is the new concentration ratio in this industry?
 - d. Suppose that after they merge, firms A and G go out of business. What is the subsequent concentration ratio in this industry?
6. The cost of pollution emanating from the chemical industry (in billions of dollars) is

$$C_P = 3P + 3P^2$$

where P is the quantity of pollutants emitted (in thousands of tons). The cost of pollution control (in billions of dollars) is

$$C_C = 7 - 5P$$

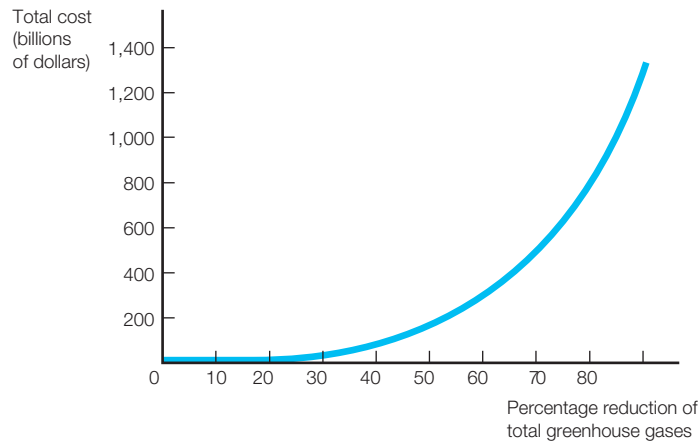
- a. What is the optimal effluent fee?
 - b. If the cost of pollution control falls by \$1 billion at each level of pollution, does this alter your answer to part (a)?
7. In the cardboard box industry, the minimum average cost is reached when a firm produces 1,000 units of output per month. At this output rate, the average cost is \$1 per unit of output. The demand curve for this product is as follows:

Price (Dollars per Unit of Output)	Quantity (Units Demanded per Month)
3.00	1,000
2.00	8,000
1.00	12,000
0.50	20,000

- a. Is this industry a natural monopoly? Why or why not?
- b. If the price is \$2, how many firms, each of which is producing output such that average cost is at a minimum, can the market support?

8. Bethlehem and Youngstown, two major steel producers, accounted for about 21% of the national steel market in the late 1950s, when they proposed to merge.
 - a. Should the two steel companies have been allowed to merge? Why or why not?
 - b. According to the companies, Bethlehem sold most of its output in the East, whereas Youngstown sold most of its output in the Midwest. Was this fact relevant? Why or why not?
 - c. The district court did not allow Bethlehem and Youngstown to merge. Yet in 1985 (as we saw in problem 1), the Department of Transportation allowed United Airlines (with about 7% of the service between Japan and the U.S. mainland) to acquire Pan Am's Pacific Division (with about 19%). How can you explain this?
9. The New York State Electric and Gas Corporation filed a request for a 10.7% increase in electric revenues. The reasons given to justify the increase were that the value of the firm's plant and equipment had increased by \$140 million, operating costs had increased, and investors required a higher rate of return.
 - a. Why should an increase in the value of the firm's plant and equipment result in an increase in the amount of revenue allowed by the Public Service Commission?
 - b. Why should an increase in operating costs have the same effect?
 - c. Why should the attitude of investors regarding what they require as a rate of return be relevant here?
10. Since early in this century, an enormous amount of attention was devoted to global warming. According to many scientists, increases in carbon dioxide and other greenhouse gases may produce significant climatic changes over the next century. To cope with this potential problem, it has been suggested that firms reduce energy consumption and switch to nonfossil fuels. William Nordhaus, a leading expert on this topic, estimated that the worldwide costs (in 1989 U.S. dollars) of various percentage reductions in the quantity of greenhouse gases emitted into the atmosphere would be as shown in the following figure.
 - a. Does this graph show the cost of pollution or the cost of pollution control?
 - b. Can this graph alone indicate the socially optimal amount of greenhouse gases that should be emitted into the atmosphere? Why or why not?
 - c. If world output is about \$20 trillion, by what percentage would that world output be reduced if the countries of the world agreed to cut greenhouse gas emission by 50%?
 - d. The single most common policy proposed to decrease greenhouse gas emissions is a carbon tax—a tax on fossil fuels in proportion to the amount of carbon they emit when burned. Why would such a tax have the desired effect?

Cost of Greenhouse Gases in the Atmosphere



Source: R. Dornbusch and J. Poterba, eds., *Global Warming: Economic Policy Responses* (Cambridge, MA: MIT Press, 1991), p. 50.

EXCEL EXERCISE: EXTERNALITIES

Suppose the following scenario. An extractive industry lies across a gorge from a town. There is a bridge across the gorge that provides the only way for the extractive industry to get its product to the marketplace. The bridge has fallen into disrepair, so the extractive industry must “light load” its trucks to cross the bridge. This light load (class 1) is significantly below the rated road weight limits that the industry’s trucks can use (class 6) to carry the product to the national market. The table below shows the extractive industry’s profits under the current weight limit (class 1) on the bridge and what the profit would be if the trucks could carry the same weight as they can on the highway system (class 6) as well as other weight classes between 1 and 6. The table also shows the costs that would have to be incurred to bring the bridge’s weight tolerance up to each of the six weight classes. The extractive industry is the only user of the bridge. The amounts in the table reflect net present values.

Weight Class	Industry’s Profit at Weight Class	Cost to Restore Bridge to the Weight Class
Class 1	100	0
Class 2	150	25
Class 3	175	40
Class 4	200	55
Class 5	225	85
Class 6	250	120

We will consider two cases. The first is that the town owns the bridge and is not required to do anything, that is, they are not required to bring the bridge up to any class standard. The extractive industry has no leverage because it can't threaten to move, as it is tied to the material in the ground at its location. In addition, it still makes economic profits (100) with the current class 1 bridge. If the two sides set down to bargain over a resolution to the "bridge problem," which class of bridge would they decide on and who would make what payments to whom?

In the second case, the extractive industry has the right to have (by law) a class 6 bridge. Now the town has no leverage. If the two sides sit down to bargain over a resolution to the "bridge problem," which class of bridge would they decide on and who would make what payments to whom?

Call up your spreadsheet. Enter 100 in cell A1, 150 in cell A2, and so on, until you've entered 250 in cell A6.

Enter 0 in cell B1, 25 in cell B2, and so on, until you've entered 120 in cell B6.

You've now duplicated the above table in your spreadsheet. As the town is required to do nothing in the first case, the extractive industry would have to (at the bare minimum) reimburse the town for any expenses the town would incur in upgrading the bridge.

Enter $=A1-B1$ in cell C1. Then click on the lower right-hand corner of cell C1 and drag down with your mouse until cell C6. Column C gives the net profits of the extractive industry for each class of bridge after reimbursing the town for its expenses in upgrading its bridge. Find the largest number in column C by observation or entering $=\text{Max}(C1:C6)$ in cell C7. This will be 145 cell C4. This is highest net profit for the extractive industry. Thus, we would expect the negotiations to lead to a class 4 bridge with the extractive industry paying (at a minimum) the town's expense of 55 to upgrade the bridge from class 1 to class 4 status.

But if the town was really good at bargaining, could it get more? Remember, the town doesn't have to do anything. That means the town could stay with the class 1 bridge relegating the industry to a profit of 100. Under the solution of having the industry pay the town the expense of upgrading the bridge to class 4, the industry increases its profit from 100 to 145. The town (which had the leverage) "breaks even." Suppose the town (which doesn't have to do anything) told the industry, "we'll bring the bridge up to class 4 for a payment of just less than 100." Would the extractive industry pay? If it did, it'd have a class 4 bridge and net profits of 200 minus a little bit less than 100, which yields a profit of a little more than 100. And that's better than the 100 it is getting now with a class 1 bridge.

So in the first case, we'd expect a class 4 bridge with a payment of no less than 55 nor more than 100 passing from the industry to the town. What the actual payment would be depends on the bargaining skills of each party. Given that the leverage lies with the town, we'd expect the payment to be closer to 100.

What about case 2? Here the industry is entitled by law to a class 6 bridge. But another way of saying this is that the industry is entitled to profits of 250. The industry doesn't care where these profits come from as long as it gets 250.

Let's solve this one.

Enter 100 in cell A8, 150 in cell A9, and so on, until you've entered 250 in cell A13.

Enter 0 in cell B8, 25 in cell B9, and so on, until you've entered 120 in cell B13.

Columns A and B duplicate the initial data table.

Enter $=A13-A8$ in cell C8, $=A13-A9$ in cell C9, $=A13-A10$ in cell C10, $=A13-A11$ in cell C11, $=A13-A12$ in cell C12, and $=A13-A13$ in cell C13. Column C represents the amount of money the town owes the industry if the town decides to improve the bridge to a class that is not class 6. For instance, if the town decides to keep the bridge at class 1, the industry will earn 100 from sales of its product. But since the industry is entitled to the profits associated with a class 6 bridge (250), the town will have to pay the industry the difference, that is, $250-100 = A13-A9$.

What are the town's total expenses for each bridge class? Enter $=B8+C8$ in cell D8 and click and drag until cell D13. Column D is the town's total cost for each bridge class, that is, the cost of upgrading the bridge plus the cost of compensating the industry for lost profit. Searching column D for the lowest number or entering $=\text{Min}(D8:D13)$ in cell D14, yields 105 in cell D11. Note that this entails a class 4 bridge just as in case 1. Note also that the town incurs a 55 cost in upgrading the bridge and a 50 payment because the industry's profit is 200 with a class 4 bridge and the industry is guaranteed 250 with the class 6 bridge.

But if the industry wanted to, it could insist that the town upgrade the bridge to class 6. That would cost the town $120 - 55 = 65$ more than to upgrade to the class 4 bridge. Going to a class 6 bridge would cost the town 65 more in bridge expense. But going to a class 6 bridge (from a class 4) would save the town 50 in terms of compensating the industry for its lost profits. Thus, if the industry wished, it could accept (at minimum) a 50 payment for its lost profits from having a class 4 bridge. Or it could hold out for slightly less than 65 (the extra cost the town would incur if the industry insisted on a class 6 bridge). Where between the 50 and 65 payment the actual payment would be depends on the bargaining skills of the parties. But since the leverage lies with the industry, we'd expect the payment to be closer to 65.

Notice that no matter who held the leverage, the optimal solution would be a class 4 bridge. Of course, the distribution of wealth differed significantly depending on who had the leverage (the property rights to having a class 1 or a class 6 bridge). This indifference solution is attributable to Coase.

CHAPTER 18

LEARNING OBJECTIVES

Functional Relationships

Marginal Analysis

Relationships among Total,
Marginal, and Average Values

The Concept of a Derivative

How to Find a Derivative

Using Derivatives to Solve
Maximization and Minimization
Problems

Marginal Cost Equals Marginal
Revenue and the Calculus of
Optimization

Partial Differentiation and the
Maximization of Multivariable
Functions

Constrained Optimization

Lagrangian Multipliers

Comparing Incremental Costs
with Incremental Revenues

Summary

Problems

OPTIMIZATION TECHNIQUES

To begin this chapter, we describe marginal analysis—a powerful tool that illuminates many central aspects of decision making. Economists think at the margins. Virtually all the rules we study about optimal behavior of firms and managers are driven by this concept.

Next, we examine the basic elements of differential calculus, including the rules of differentiation and the use of a derivative to maximize a function such as profit or minimize one such as cost. Differentiation tells us what changes will occur in one variable (the dependent variable) when a small (marginal) change is made in another variable (the independent variable). Therefore, marginal analysis can be implemented by the use of differentiation.

Finally, we examine constrained optimization, including Lagrangian multipliers. When managers want to maximize profit, such maximization or minimization is often subject to constraints (such as producing a certain output to adhere to a contract or utilizing a certain amount of labor in a union agreement).

FUNCTIONAL RELATIONSHIPS

Frequently a relationship between economic variables is represented by a table or graph. Although tables and graphs are helpful, another way of expressing economic relationships is with equations. For example, how can the relationship

between the number of units sold and the price be expressed in an equation? One way is to use the following functional notation

$$Q = f(P) \quad (18.1)$$

where Q is the number of units sold and P is price. This equation is read as “The number of units sold is a function of price,” which means the number of units sold *depends* on price. In other words, the number of units sold is the *dependent* variable, and price is the *independent* variable.

Equation (18.1) is useful, but it does not tell us *how* the number of units sold depends on price. A more specific representation of this relationship is

$$Q = 200 - 5P \quad (18.2)$$

This equation says that if the price equals \$10, the number of units sold should be $200 - 5(10) = 150$.

MARGINAL ANALYSIS

The **marginal value** of a dependent variable is defined as the change in this dependent variable associated with a one-unit change in a particular independent variable. As an illustration, consider Table 18.1, which shows in columns 1 and 2 the total profit of the Roland Corporation if we vary the number of units produced. In this case, total profit is the dependent variable and output is the independent variable. Therefore, the marginal value of profit, called the **marginal profit**, is the change in total profit associated with a one-unit change in output.

Column 3 of Table 18.1 shows the value of marginal profit. If output increases from zero to one unit, column 2 shows the total profit increases by \$100 (from \$0 to \$100). Therefore, the marginal profit in column 3 equals \$100 if the output is one unit. If the output increases from one to two units, the total profit increases by \$150 (from \$100 to \$250). Therefore, the marginal profit in column 3 equals \$150 if the output is increased from one to two units.

The central point about a marginal relationship of this sort is that the dependent variable—in this case, total profit—is maximized when its marginal value shifts from positive to negative. To see this, consider Table 18.1. So long as marginal profit is positive, the Roland Corporation can raise its total profit by increasing output. For example, if output increases from five to six units, the marginal profit is positive (\$150); therefore, the firm’s total profit goes up (by \$150). But when marginal profit shifts from positive to negative, total profit *falls* with any further increase in output. In Table 18.1, this point is reached when the firm produces seven units of output. If output increases beyond seven units, marginal profit shifts from positive to negative—and total profit goes down (by \$50). So we can see that the

Marginal value The change in the dependent variable associated with a one-unit change in a particular independent variable.

Marginal profit The change in total profit associated with a one-unit change in output.

TABLE 18.1

Relationship between Output and Profit: Roland Corporation

(1) Number of Units of Output per Day	(2) Total Profit	(3) Marginal Profit	(4) Average Profit
0	0	—	—
1	100	100	100
2	250	150	125
3	600	350	200
4	1,000	400	250
5	1,350	350	270
6	1,500	150	250
7	1,550	50	221.4
8	1,500	−50	187.5
9	1,400	−100	155.5
10	1,200	−200	120

dependent variable—in this case, total profit—is maximized when its marginal value shifts from positive to negative.

Because managers are interested in determining how to maximize profit (or other performance measures), this is a useful result. It emphasizes the importance of looking at marginal values—and the hazards that may arise if average values are used instead. In Table 18.1, **average profit**—that is, total profit divided by output—is shown in column 4. It may seem reasonable to choose the output level that gives the highest average profit; countless managers have done so. But this is not the correct decision if managers want to maximize profit. Instead, as stressed in the previous paragraph, managers should choose the output level at which marginal profit shifts from positive to negative.

To prove this, we need only find the output level in Table 18.1 at which average profit is highest. Based on a comparison of the figures in column 4, this output level is five units; and according to column 2, the total profit at this output level equals \$1,350. But we found that the output level at which marginal profit shifts from positive to negative is seven units; and according to column 2, the total profit at this point equals \$1,550. In other words, the total profit is \$200 higher if the output level is seven rather than five units. Thus if managers of this firm were to choose the output level at which average profit is highest, they would sacrifice \$200 per day in profits.

It is important to understand the relationship between average and marginal values. Because the marginal value represents the change in the total, the average

Average profit The total profit divided by output.

value must increase if the marginal value is greater than the average value. Similarly, the average value must decrease if the marginal value is less than the average value. Table 18.1 illustrates these propositions. For the first to fifth units of output, the marginal profit is greater than the average profit. Because the extra profit from each additional unit is greater than the average, the average is pulled up as more units are produced. For the sixth to tenth units of output, the marginal profit is less than the average profit. Because the extra profit from each additional unit is less than the average, the average value is pulled down as more units are produced.

RELATIONSHIPS AMONG TOTAL, MARGINAL, AND AVERAGE VALUES

To further explore the relationships among total, marginal, and average values, consider Figure 18.1, which shows the relationships among total, average, and marginal profit, on the one hand, and output, on the other hand, for the Roland Corporation. The relationship between output and profit is exactly the same as in Table 18.1; but rather than using particular numbers to designate output or profit, we use symbols such as Q_0 and Q_1 for output levels and π_0 for a profit level. This makes the results valid in general, not just for a particular set of numerical values.

Note that Figure 18.1 contains two panels. The upper panel (panel A) shows the relationship between total profit and output levels, whereas the lower panel (panel B) shows the relationship between average profit and marginal profit, on the one hand, and output levels, on the other. The horizontal scale of panel A is the same as that of panel B, so a given output level, like Q_0 , is the same distance from the origin (along the horizontal axis) in panel A as in panel B.

In practice we seldom are presented with data concerning both (1) the relationship between total profit and output and (2) the relationship between average profit and output because it is relatively simple to derive the latter relationship from the former. How can this be done? Take any output level, say Q_0 . *At this output level, the average profit equals the slope of the straight line from the origin to point E, the point on the total profit curve corresponding to output level Q_0 .* To see that this is the case, note that the average profit at this output level equals π_0/Q_0 , where π_0 is the level of total profit if the output level is Q_0 . Because the slope of any straight line equals the vertical distance between two points on the line divided by the horizontal distance between them, the slope of the line from the origin to point E equals π_0/Q_0 .¹ Thus the slope of line OE equals the average profit at this output level. (In other words, K_0 in panel B of Figure 18.1 is equal to the slope of line OE.) To determine the relationship between average profit and output from the relationship between total profit and output, we repeat this procedure for each level of output, not just Q_0 . The resulting average profit curve is shown in panel B.

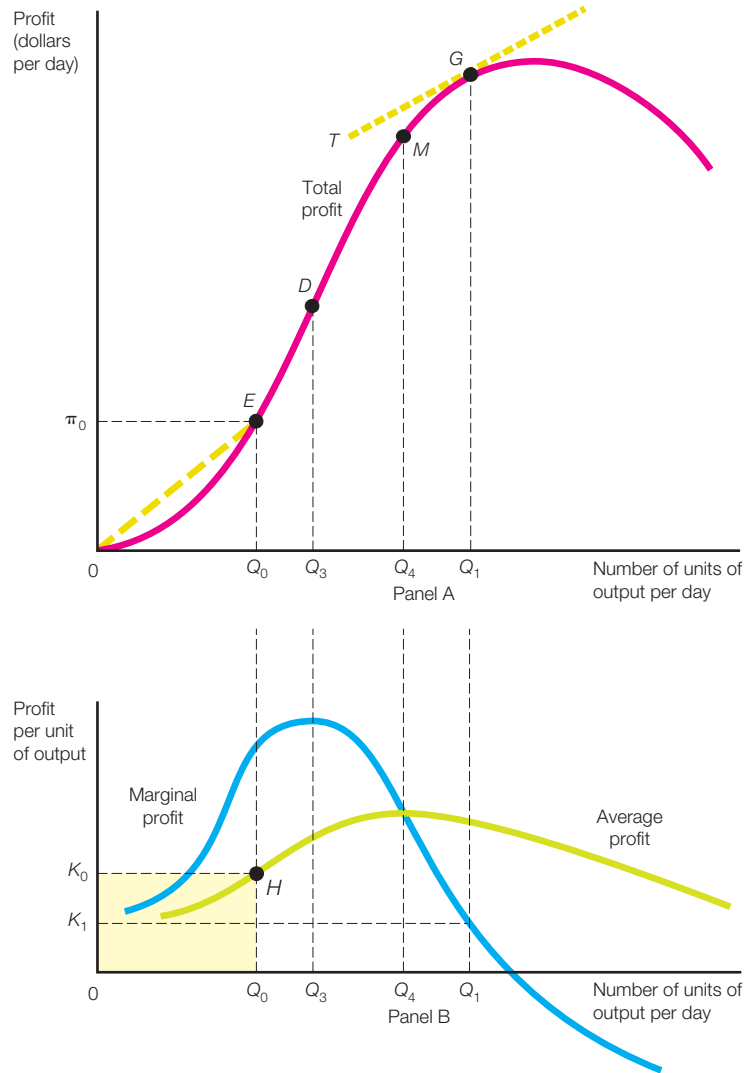
Turning to the relationship between marginal profit and output (in panel B), it is relatively simple to derive this relationship also from the relationship between

1. The vertical distance between the origin and the point E equals π_0 , and the horizontal distance between these two points equals Q_0 . Therefore, the vertical distance divided by the horizontal distance equals π_0/Q_0 .

FIGURE 18.1

Total Profit, Average Profit, and Marginal Profit: Roland Corporation

The average and marginal profit curves in panel B can be derived geometrically from the total profit curve in panel A.



total profit and output (in panel A). Take any output level, say Q_1 . At this output level, the marginal profit equals the slope of the tangent to the total profit curve (in panel A) at the point where the output level is Q_1 . In other words, the marginal profit equals the slope of line T in Figure 18.1, which is tangent to the total profit curve at point G . As a first step toward seeing why this is true, consider Figure 18.2,

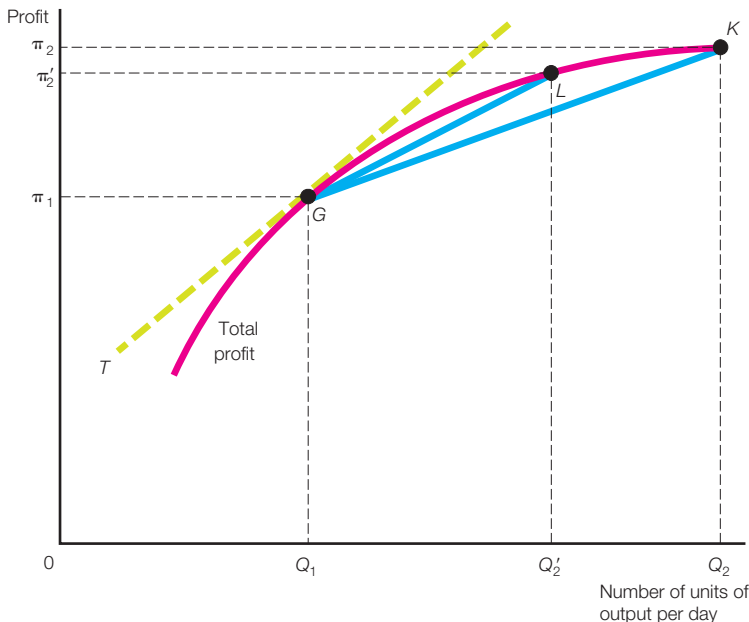
which provides a magnified picture of the total profit curve in the neighborhood of point G.

Recall that marginal profit is defined as the extra profit resulting from a very small increase (specifically, a one-unit increase) in output. If the output level increases from Q_1 to Q_2 , the total profit increases from π_1 to π_2 , as shown in Figure 18.2. Therefore, the extra profit per unit of output is $(\pi_2 - \pi_1)/(Q_2 - Q_1)$, which is the slope of the GK line. But this increase in output is rather large. Suppose we decrease Q_2 so it is closer to Q_1 . In particular, let the new value of Q_2 be Q'_2 . If output increases from Q_1 to Q'_2 , the extra profit per unit of output equals $(\pi'_2 - \pi_1)/(Q'_2 - Q_1)$, which is the slope of the GL line. If we further decrease Q_2 until the distance between Q_1 and Q_2 is extremely small, the slope of the tangent (line T) at point G becomes a good estimate of $(\pi_2 - \pi_1)/(Q_2 - Q_1)$. In the limit, for changes in output in a very small neighborhood around Q_1 , the slope of the tangent is marginal profit. (This slope equals K_1 in panel B of Figure 18.1.) To determine the relationship between marginal profit and output from the relationship between total profit and output, we repeat this procedure for each level

FIGURE 18.2

Marginal Profit Equals the Slope of the Tangent to the Total Profit Curve

As the distance between Q_1 and Q_2 becomes extremely small, the slope of line T becomes a good estimate of $(\pi_2 - \pi_1)/(Q_2 - Q_1)$.



of output, not just Q_1 . The resulting marginal profit curve is shown in panel B of Figure 18.1.

Sometimes we are given an average profit curve like that in panel B of Figure 18.1 but not the total profit curve. To derive the latter curve from the former, note that total profit equals average profit times output. Hence, if output equals Q_0 , total profit equals K_0 times Q_0 . In other words, π_0 in panel A equals the area of rectangle OK_0HQ_0 in panel B. To derive the relationship between total profit and output from the relationship between average profit and output, we repeat this procedure for each level of output. That is, we find the area of the appropriate rectangle of this sort corresponding to each output level, not just Q_0 . The resulting total profit curve is shown in panel A.

Finally, two further points should be made concerning the total, average, and marginal profit curves in Figure 18.1. First, you should be able to tell by a glance at panel A that the marginal profit increases as the output level rises from zero to Q_3 and that it decreases as output rises further. Why is this so obvious from panel A? Because the slope of the total profit curve increases as we move from the origin to point D . In other words, lines drawn tangent to the total profit curve become steeper as we move from the origin to point D . Because marginal profit equals the slope of this tangent, it must increase as the output level rises from zero to Q_3 . To the right of point D , the slope of the total profit curve decreases as the output level increases. That is, lines drawn tangent to the total profit curve become less steep as we move to the right of point D . Consequently, because marginal profit equals the slope of this tangent, it too must decrease when the output level rises beyond Q_3 .

Second, panel B of Figure 18.1 confirms the following proposition: *The average profit curve must be rising if it is below the marginal profit curve, and it must be falling if it is above the marginal profit curve.* At output levels below Q_4 , the average profit curve is below the marginal profit curve; therefore, the average profit curve is rising because the higher marginal profits are pulling up the average profits. At output levels above Q_4 , the average profit curve is above the marginal profit curve; therefore, the average profit curve is falling because the lower marginal profits are pulling down the average profits. At Q_4 , the straight line drawn from the origin to point M is just tangent to the total cost curve. Therefore, the average profit and marginal profit are equal at output level Q_4 .

THE CONCEPT OF A DERIVATIVE

In the case of the Roland Corporation, we used Table 18.1 (which shows the relationship between the firm's output and profit) to find the profit-maximizing output level. Frequently a table of this sort is too cumbersome or inaccurate to be useful for this purpose. Instead we use an equation to represent the relationship between the variable we are trying to maximize (in this case, profit) and the variable or variables under the control of the decision maker (in this case, output).

Given an equation of this sort, we can employ the powerful concepts and techniques of differential calculus to find optimal decision solutions.

In previous sections, we defined the *marginal value* as the change in a dependent variable resulting from a one-unit change in an independent variable. If Y is the dependent variable and X is the independent variable

$$Y = f(X) \quad (18.3)$$

according to the notation in equation (18.1). Using Δ (called *delta*) to denote change, we can express a change in the independent variable as ΔX , and we can express a change in the dependent variable as ΔY . Thus the marginal value of Y is estimated by

$$\frac{\text{Change in } Y}{\text{Change in } X} = \frac{\Delta Y}{\Delta X} \quad (18.4)$$

For example, if a two-unit increase in X results in a one-unit increase in Y , $\Delta X = 2$ and $\Delta Y = 1$, then the marginal value of Y is about one-half. That is, the dependent variable Y increases by about one-half if the independent variable X increases by one.²

Unless the relationship between Y and X can be represented as a straight line (as in Figure 18.3), the value of $\Delta Y/\Delta X$ is not constant. For example, consider the relationship between Y and X in Figure 18.4. If a movement occurs from point G to point H , a relatively small change in X (from X_1 to X_2) is associated with a big change in Y (from Y_1 to Y_2). Therefore, between points G and H , the value of $\Delta Y/\Delta X$, which equals $(Y_2 - Y_1)/(X_2 - X_1)$, is relatively large. On the other hand, if a movement occurs from point K to point L , a relatively large change in X (from X_3 to X_4) is associated with a small change in Y (from Y_3 to Y_4). Consequently, between points K and L , the value of $\Delta Y/\Delta X$, which equals $(Y_4 - Y_3)/(X_4 - X_3)$, is relatively small.

The value of $\Delta Y/\Delta X$ is related to the steepness or flatness of the curve in Figure 18.4. Between points G and H the curve is relatively *steep*; this means a *small* change in X results in a *large* change in Y . Consequently $\Delta Y/\Delta X$ is relatively large. Between points K and L the curve is relatively *flat*; this means a *large* change in X results in a *small* change in Y . Consequently $\Delta Y/\Delta X$ is relatively small.

The derivative of Y with respect to X is defined as the limit of $\Delta Y/\Delta X$ as ΔX approaches zero. Because the derivative of Y with respect to X is denoted by dY/dX , this definition can be restated as

$$\frac{dY}{dX} = \lim_{\Delta X \rightarrow 0} \frac{\Delta Y}{\Delta X} \quad (18.5)$$

which is read “The derivative of Y with respect to X equals the limit of the ratio $\Delta Y/\Delta X$ as ΔX approaches zero.” To understand what is meant by a limit, consider the function $(X - 2)$. What is the limit of this function as X approaches 2? Clearly,

2. Why do we say that Y increases by about $\frac{1}{2}$ rather than by exactly $\frac{1}{2}$? Because Y may not be linearly related to X .

FIGURE 18.3

Linear Relationships between Y and X

The relationship between Y and X can be represented as a straight line.

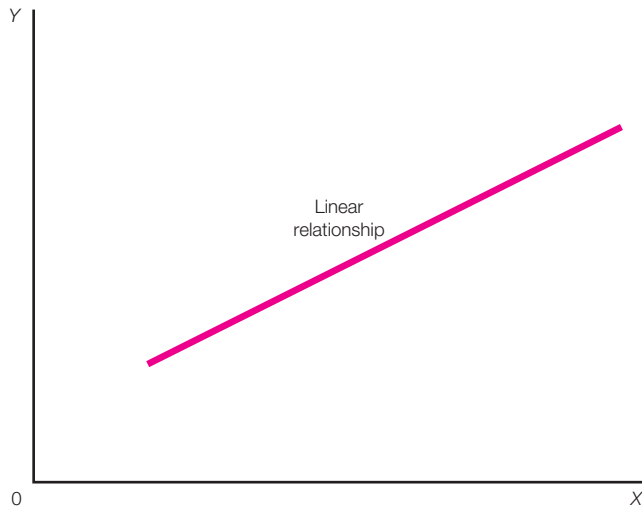
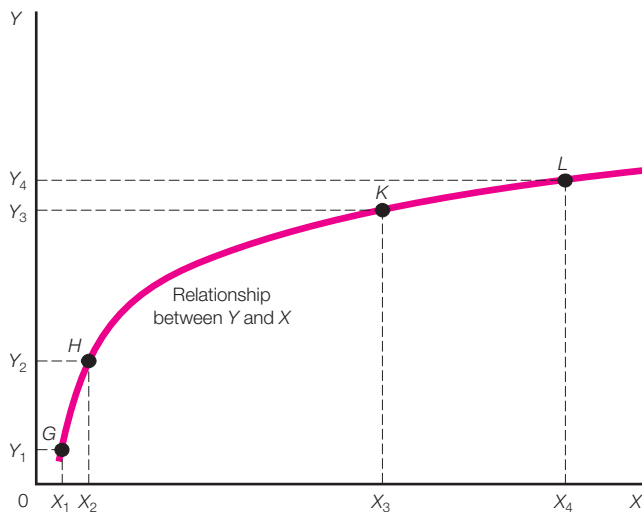


FIGURE 18.4

How the Value of $\Delta Y/\Delta X$ Varies Depending on the Steepness or Flatness of the Relationship between Y and X

Between points G and H , the curve is steep, so $\Delta Y/\Delta X$ is large. Between points K and L , the curve is flat, so $\Delta Y/\Delta X$ is small.



as X gets closer and closer to 2, $(X - 2)$ gets closer and closer to zero. What is the limit of this function as X approaches zero? Clearly, as X gets closer and closer to zero, $(X - 2)$ gets closer and closer to -2 .

Graphically, the derivative of Y with respect to X equals the *slope* of the curve showing Y (on the vertical axis) as a function of X (on the horizontal axis). To see this, suppose we want to find the value of the derivative of Y with respect to X when X equals X_5 in Figure 18.5. A rough measure is the value of $\Delta Y/\Delta X$ when a movement is made from point A to point C ; this measure equals

$$(Y_7 - Y_5)/(X_7 - X_5)$$

which is the slope of the AC line. A better measure is the value of $\Delta Y/\Delta X$ when a movement is made from point A to point B ; this measure equals

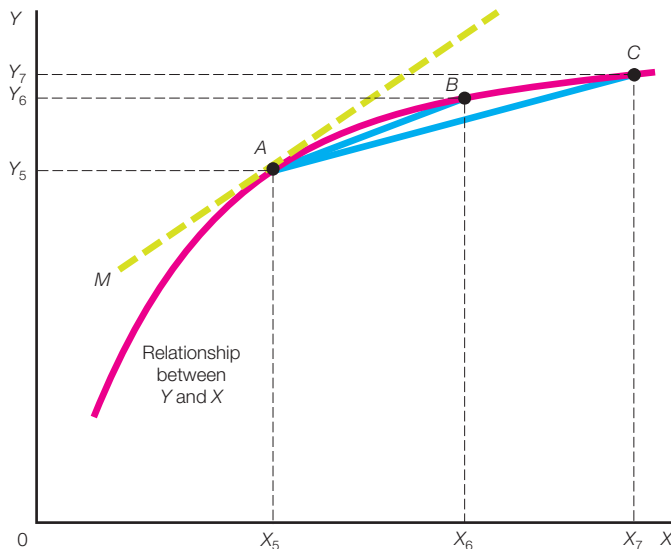
$$(Y_6 - Y_5)/(X_6 - X_5)$$

which is the slope of the AB line. Why is the latter measure better than the former? Because the distance between points A and B is less than the distance between points A and C , and what we want is the value of $\Delta Y/\Delta X$ when ΔX is as small as possible. *Clearly, in the limit, as ΔX approaches zero, the ratio $\Delta Y/\Delta X$ is equal to the slope of the line M , which is drawn tangent to the curve at point A .*

FIGURE 18.5

Derivative as the Slope of the Curve

When X equals X_5 , the derivative of Y with respect to X equals the slope of line M , the tangent to the curve at point A .



HOW TO FIND A DERIVATIVE

Managers want to know how to optimize performance. If Y is some measure of organizational performance and X is a variable under a particular manager's control, he or she would like to know the value of X that maximizes Y . To find out this value, managers want to approximate the derivative of Y with respect to X . In this section, we learn how to find this derivative.

Derivatives of Constants

If the dependent variable Y is a constant, its derivative with respect to X is always zero. That is, if $Y = a$ (where a is a constant)

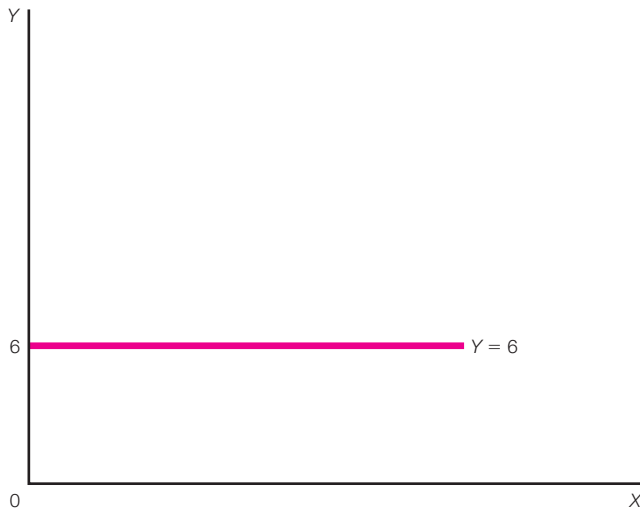
$$\frac{dY}{dX} = 0 \quad (18.6)$$

EXAMPLE Suppose $Y = 6$, as shown in Figure 18.6. Because the value of Y does not change as X varies, dY/dX is equal to zero. To see how this is shown geometrically, recall from the previous section that dY/dX equals the slope of the curve showing Y as a function of X . As is evident from Figure 18.6, this slope equals zero (since it is a horizontal line), which means dY/dX is equal to zero.

FIGURE 18.6

Case in Which $Y = 6$

In this case, dY/dX equals zero because the slope of this horizontal line equals zero.



Derivatives of Power Functions

A power function can be expressed as

$$Y = aX^b$$

where a and b are constants. If the relationship between X and Y is of this kind, the derivative of Y with respect to X equals b times a multiplied by X raised to the $(b - 1)$ power

$$\frac{dY}{dX} = baX^{b-1} \quad (18.7)$$

EXAMPLE Suppose $Y = 3X$, which is graphed in panel A of Figure 18.7. Applying equation (18.7), we find that

$$\frac{dY}{dX} = 1 \times 3 \times X^0 = 3$$

because $a = 3$ and $b = 1$. Therefore, the value of dY/dX graphed in panel B of Figure 18.7 is 3, regardless of the value of X . This makes sense: The slope of the line in panel A is 3, regardless of the value of X . Recall once again from the previous section that dY/dX equals the slope of the curve showing Y as a function of X . In this case (as in Figure 18.6), the “curve” is a straight line.

EXAMPLE Suppose $Y = 2X^2$, which is graphed in panel A of Figure 18.8. Applying equation (18.7), we find that

$$\frac{dY}{dX} = 2 \times 2 \times X^1 = 4X$$

because $a = 2$ and $b = 2$. Therefore, the value of dY/dX , which is graphed in panel B of Figure 18.8, is proportional to X . As we would expect, dY/dX is negative when the slope of the curve in panel A is negative and positive when this slope is positive. Why? Because, as we have stressed repeatedly, dY/dX equals this slope.

We make a plea here for the actual ease of using calculus. All the calculus necessary to teach the concepts of managerial economics in this book can be summed up in one easy-to-remember formula.

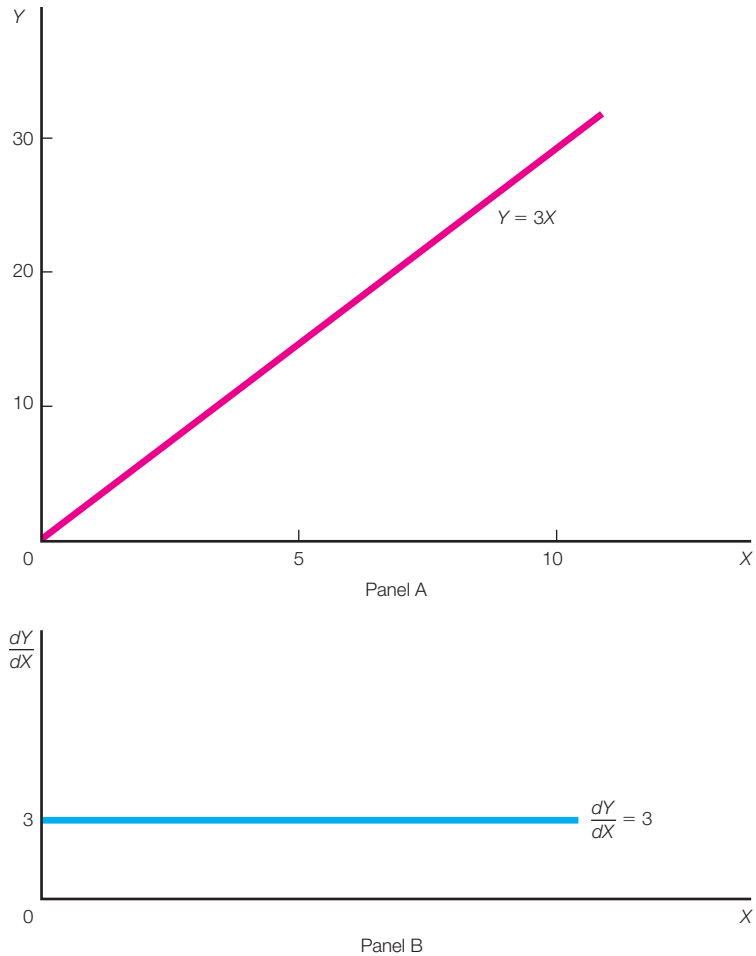
If $y = kx^n$, where y is the dependent variable, that is, its value depends on the value of x , the independent variable. The value of X determines y 's value. n and k are parameters (i.e., numbers). If you want to find the impact a change in the independent variable x has on the dependent variable y , you differentiate y with respect to x , that is, dy/dx . Now here's the easy-to-remember formula

$$dy/dx = nkx^{n-1}$$

FIGURE 18.7

Case in Which $Y = 3X$

In this case, dY/dX equals 3 because the slope of the line in panel A equals 3.



So if $y = 4x^3$, that is, $n = 3$ and $k = 4$, then $dy/dx = 12x^2$ with $nk = 3 \times 4 = 12$ and $n - 1 = 2$. Memorize the formula above, and you understand all the calculus to answer every numerical example in this book.

Derivatives of Sums and Differences

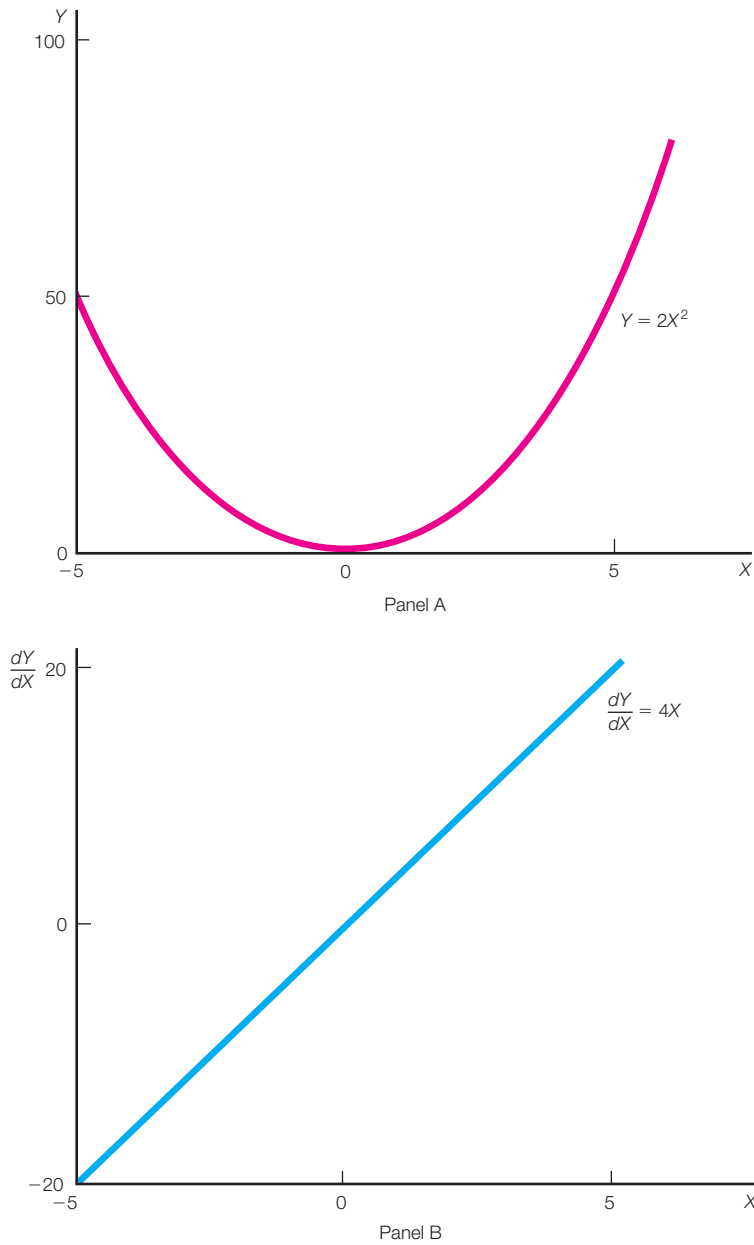
Suppose U and W are two variables, each of which depends on X . That is

$$U = g(X) \text{ and } W = h(X)$$

FIGURE 18.8

Case in Which $Y = 2X^2$

In this case, $dY/dX = 4X$ because the slope of the curve in panel A equals $4X$.



The functional relationship between U and X is denoted by g , and that between W and X is denoted by h . Suppose further that

$$Y = U + W$$

In other words, Y is the sum of U and W . If so, the derivative of Y with respect to X equals the sum of the derivatives of the individual terms

$$\frac{dY}{dX} = \frac{dU}{dX} + \frac{dW}{dX} \quad (18.8)$$

On the other hand, if

$$Y = U - W$$

the derivative of Y with respect to X equals the difference between the derivatives of the individual terms

$$\frac{dY}{dX} = \frac{dU}{dX} - \frac{dW}{dX} \quad (18.9)$$

EXAMPLE Consider the case in which $U = g(X) = 3X^3$ and $W = h(X) = 4X^2$. If $Y = U + W = 3X^3 + 4X^2$

$$\frac{dY}{dX} = 9X^2 + 8X \quad (18.10)$$

To see why, recall from equation (18.8) that

$$\frac{dY}{dX} = \frac{dU}{dX} + \frac{dW}{dX} \quad (18.11)$$

Applying equation (18.7), we have

$$\frac{dU}{dX} = 9X^2 \text{ and } \frac{dW}{dX} = 8X$$

Substituting these values of the derivatives into equation (18.11), we obtain equation (18.10).

EXAMPLE Suppose $Y = U - W$, where $U = 8X^2$ and $W = 9X$. Then

$$\frac{dY}{dX} = 16X - 9$$

because, according to equation (18.9)

$$\frac{dY}{dX} = \frac{dU}{dX} - \frac{dW}{dX}$$

and, applying equation (18.7), we have

$$\frac{dU}{dX} = 16X \text{ and } \frac{dW}{dX} = 9$$

Derivatives of Products

The derivative of the product of two terms is equal to the sum of the first term multiplied by the derivative of the second plus the second term multiplied by the derivative of the first. Consequently, if $Y = UW$, we have

$$\frac{dY}{dX} = U \frac{dW}{dX} + W \frac{dU}{dX} \quad (18.12)$$

EXAMPLE If $Y = 6X(3 - X^2)$, we can let $U = 6X$ and $W = 3 - X^2$; then

$$\begin{aligned} \frac{dY}{dX} &= 6X \frac{dW}{dX} + (3 - X^2) \frac{dU}{dX} \\ &= 6X(-2X) + (3 - X^2)(6) \\ &= -12X^2 + 18 - 6X^2 \\ &= 18 - 18X^2 \end{aligned}$$

The first term, $6X$, is multiplied by the derivative of the second term, $-2X$, and the result is added to the second term, $3 - X^2$, multiplied by the derivative of the first, 6 . As indicated, the result is $18 - 18X^2$.

Derivatives of Quotients

If $Y = U/W$, the derivative of Y with respect to X equals

$$\frac{dY}{dX} = \frac{W(dU/dX) - U(dW/dX)}{W^2} \quad (18.13)$$

In other words, the derivative of the quotient of two terms equals the denominator times the derivative of the numerator minus the numerator times the derivative of the denominator—all divided by the square of the denominator.

EXAMPLE Consider the problem of finding the derivative of the expression

$$Y = \frac{5X^3}{3 - 4X}$$

If we let $U = 5X^3$ and $W = 3 - 4X$

$$\frac{dU}{dX} = 15X^2 \text{ and } \frac{dW}{dX} = -4$$

Consequently, applying equation (18.13), we have

$$\begin{aligned}\frac{dY}{dX} &= \frac{(3 - 4X)(15X^2) - 5X^3(-4)}{(3 - 4X)^2} \\ &= \frac{45X^2 - 60X^3 + 20X^3}{(3 - 4X)^2} \\ &= \frac{45X^2 - 40X^3}{(3 - 4X)^2}\end{aligned}$$

Derivatives of a Function of a Function (the Chain Rule)

Sometimes a variable depends on another variable, which in turn depends on a third variable. For example, suppose $Y = f(W)$ and $W = g(X)$. Under these circumstances, the derivative of Y with respect to X equals

$$\frac{dY}{dX} = \left(\frac{dY}{dW}\right)\left(\frac{dW}{dX}\right) \quad (18.14)$$

In other words, to find this derivative, we find the derivative of Y with respect to W and multiply it by the derivative of W with respect to X .

EXAMPLE Suppose $Y = 4W + W^3$ and $W = 3X^2$. To find dY/dX , we begin by finding dY/dW and dW/dX

$$\begin{aligned}\frac{dY}{dW} &= 4 + 3W^2 \\ &= 4 + 3(3X^2)^2 \\ &= 4 + 27X^4 \\ \frac{dW}{dX} &= 6X\end{aligned}$$

Then, to find dY/dX , we multiply dY/dW and dW/dX

$$\begin{aligned}\frac{dY}{dX} &= (4 + 27X^4)(6X) \\ &= 24X + 162X^5\end{aligned}$$

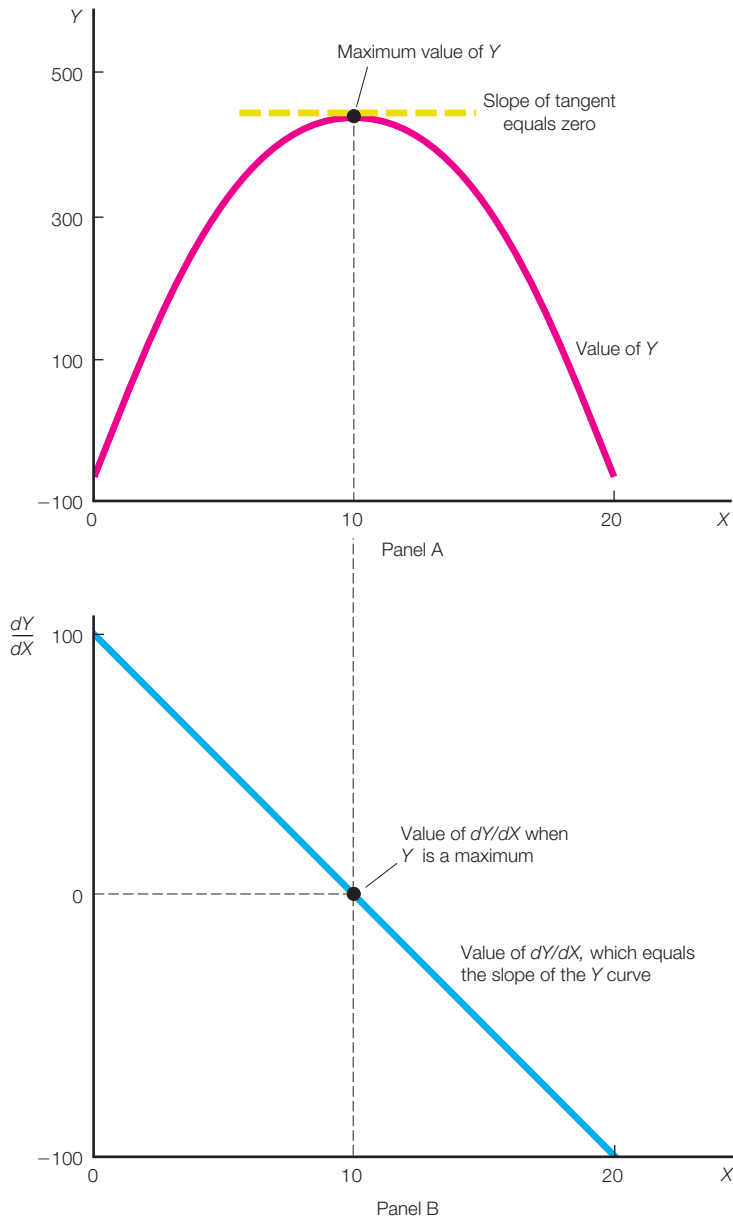
USING DERIVATIVES TO SOLVE MAXIMIZATION AND MINIMIZATION PROBLEMS

Having determined how to find the derivative of Y with respect to X , we now see how to determine the value of X that maximizes or minimizes Y . *The central point is that a maximum or minimum point can occur only if the slope of the curve showing Y on the vertical axis and X on the horizontal axis equals zero.* To see this, suppose Y equals the profit of the Monroe Company and X is its output level. If the relationship between Y and X is as shown by the curve in panel A of Figure 18.9, the

FIGURE 18.9

Value of the Derivative When Y Is a Maximum

When Y is a maximum (at $X = 10$), dY/dX equals zero.



maximum value of Y occurs when $X = 10$, and at this value of X the slope of the curve equals zero.

Because the derivative of Y with respect to X equals the slope of this curve, it follows that Y is a maximum or minimum only if this derivative equals zero. To see that Y really is maximized when this derivative equals zero, note the relationship between Y and X in Figure 18.9 is

$$Y = -50 + 100X - 5X^2 \quad (18.15)$$

which means that

$$\frac{dY}{dX} = 100 - 10X \quad (18.16)$$

Therefore, if this derivative equals zero

$$\begin{aligned} 100 - 10X &= 0 \\ X &= 10 \end{aligned}$$

This is the value of X where Y is maximized. The key point here is that *to find the value of X that maximizes or minimizes Y , we must find the value of X where this derivative equals zero*. Panel B of Figure 18.9 shows graphically that this derivative equals zero when Y is maximized.

While we know the derivative equals 0 at $X = 10$, we must distinguish between a point on the curve where Y is maximized and a point where Y is minimized. For example, in Figure 18.10, this derivative is zero both when $X = 5$ and when $X = 15$. When $X = 15$, Y is a maximum; when $X = 5$, Y is a minimum. To distinguish between a maximum and a minimum, we must find the *second derivative of Y with respect to X , which is denoted d^2Y/dX^2 and is the derivative of dY/dX* . For example, in Figure 18.9, the second derivative of Y with respect to X is the derivative of the function in equation (18.16); therefore, it equals -10 .

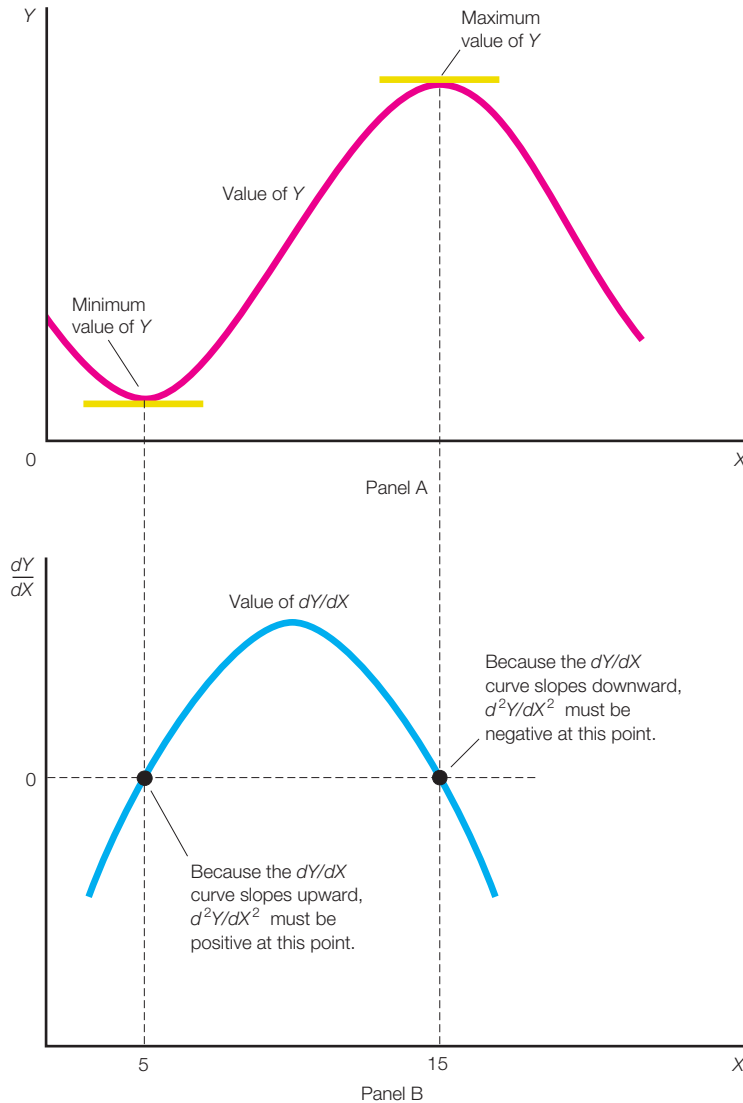
The second derivative measures the slope of the curve showing the relationship between dY/dX (the first derivative) and X . Just as the first derivative (that is, dY/dX) measures the slope of the Y curve in panel A of Figure 18.10, the second derivative (that is, d^2Y/dX^2) measures the slope of the dY/dX curve in panel B of Figure 18.10. In other words, just as the first derivative measures the slope of the total profit curve, the second derivative measures the slope of the marginal profit curve. The second derivative is important because it is always *negative* at a point of *maximization* and always *positive* at a point of *minimization*. Therefore, *to distinguish between maximization and minimization points, all we have to do is determine whether the second derivative at each point is positive or negative*.

To understand why the second derivative is always negative at a maximization point and always positive at a minimization point, consider Figure 18.10. When the second derivative is negative, this means the slope of the dY/dX curve in panel B is negative. Because dY/dX equals the slope of the Y curve in panel A,

FIGURE 18.10

Using the Second Derivative to Distinguish Maxima from Minima

At maxima (such as $X = 15$), d^2Y/dX^2 is negative; at minima (such as $X = 5$), d^2Y/dX^2 is positive.



this in turn means the slope of the Y curve decreases as X increases. At a maximum point, such as when $X = 15$, this must be the case. On the other hand, when the second derivative is positive, this means the slope of the dY/dX curve in panel B is positive, which is another way of saying that the slope of the Y curve in

panel A increases as X increases. At a minimum point, such as when $X = 5$, this must be the case.

EXAMPLE To illustrate how we can use derivatives to solve maximization and minimization problems, suppose the relationship between profit and output at the Kantor Corporation is

$$Y = -1 + 9X - 6X^2 + X^3$$

where Y equals annual profit (in millions of dollars) and X equals annual output (in millions of units). This equation is valid only for values of X that equal 3 or less; capacity limitations prevent the firm from producing more than 3 million units per year. To find the values of output that maximize or minimize profit, we find the derivative of Y with respect to X and set it equal to zero

$$\frac{dY}{dX} = 9 - 12X + 3X^2 = 0 \quad (18.17)$$

Solving this equation for X , we find that two values of X —1 and 3—result in this derivative being zero.³

To determine whether each of these two output levels maximizes or minimizes profit, we find the value of the second derivative at these two values of X . Taking the derivative of dY/dX , which is shown in equation (18.17) to equal $9 - 12X + 3X^2$, we find that

$$\frac{d^2Y}{dX^2} = -12 + 6X$$

If $X = 1$

$$\frac{d^2Y}{dX^2} = -12 + 6(1) = -6$$

Because the second derivative is negative, profit is a maximum (at 3) when output equals 1 million units. If $X = 3$

$$\frac{d^2Y}{dX^2} = -12 + 6(3) = 6$$

Because the second derivative is positive, profit is a minimum (at -1) when output equals 3 million units.

3. If an equation is of the general quadratic form, $Y = aX^2 + bX + c$, the values of X at which Y is 0 are

$$X = \frac{-b \pm (b^2 - 4ac)^{0.5}}{2a}$$

In the equation in the text, $a = 3$, $b = -12$, and $c = 9$. Hence

$$\begin{aligned} X &= \frac{12 \pm (144 - 108)^{0.5}}{6} \\ &= 2 \pm 1 \end{aligned}$$

Therefore, $dY/dX = 0$ when X equals 1 or 3.

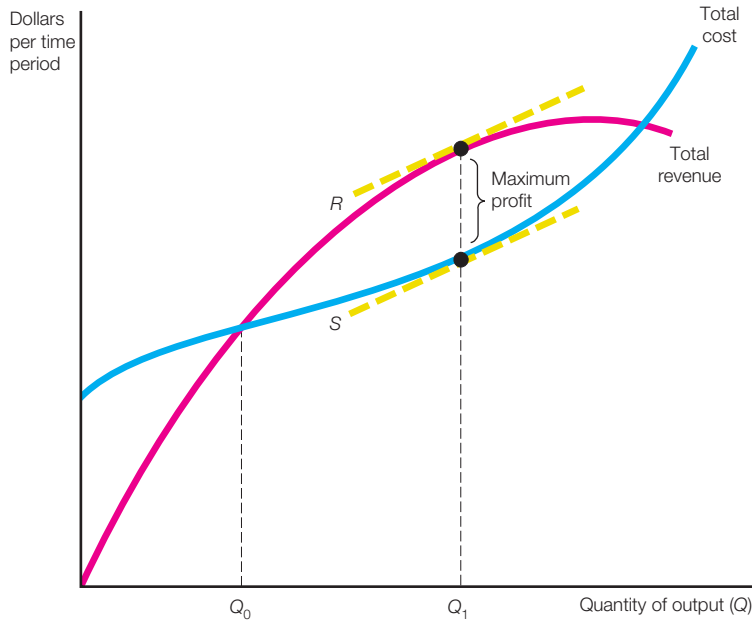
MARGINAL COST EQUALS MARGINAL REVENUE AND THE CALCULUS OF OPTIMIZATION

Once you know how elementary calculus is used to solve optimization problems, it is easy to see that the fundamental rule for profit maximization—set marginal

FIGURE 18.11

Marginal Revenue Equals Marginal Cost Rule for Profit

At the profit-maximizing output of Q_1 , marginal revenue (equal to the slope of line R) equals marginal cost (the slope of line S).



cost equal to marginal revenue—is based on the calculus of optimization. Figure 18.11 shows a firm's total cost and total revenue functions. Because total profit equals total revenue minus total cost, it equals the vertical distance between the total revenue and total cost curves at any level of output. This distance is maximized at output Q_1 , where the slopes of the total revenue and total cost curves are equal. Because the slope of the total revenue curve is marginal revenue and the slope of the total cost curve is marginal cost, profit is maximized when marginal cost equals marginal revenue.

Inspection of Figure 18.11 shows that Q_1 must be the profit-maximizing output. Output levels below Q_0 result in losses (because total cost exceeds total revenue) and obviously do not maximize profit. As output increases beyond Q_0 , total revenue rises more rapidly than total cost, so profit must be going up. So long as the slope of the total revenue curve (which equals marginal revenue) exceeds the slope of the total cost curve (which equals marginal cost), profit will continue to rise as output increases. But when these slopes become equal (that is, when marginal revenue equals marginal cost), profit no longer will rise but will be at

a maximum. These slopes become equal at an output level of Q_1 , and so this must be the profit-maximizing output level. After output Q_1 , profit decreases because marginal cost exceeds marginal revenue.

Using calculus, we can readily understand why managers maximize profit by setting marginal cost equal to marginal revenue. The first thing to note is that

$$\pi = TR - TC$$

where π equals total profit, TR equals total revenue, and TC equals total cost. Taking the derivative of π with respect to Q (output), we find that

$$\frac{d\pi}{dQ} = \frac{dTR}{dQ} - \frac{dTC}{dQ}$$

For π to be a maximum, this derivative must be zero, so it must be true that

$$\frac{dTR}{dQ} = \frac{dTC}{dQ} \quad (18.18)$$

And because marginal revenue is defined as dTR/dQ and marginal cost is defined as dTC/dQ , marginal revenue must equal marginal cost.⁴

PARTIAL DIFFERENTIATION AND THE MAXIMIZATION OF MULTIVARIABLE FUNCTIONS

Up to this point, we have examined situations in which a variable depends on only one other variable. Although such situations exist, in many cases a variable depends on a number (often a large number) of other variables, not just one. For example, the Merrimack Company produces two goods, and its profit depends on the amount that it produces of each good. That is

$$\pi = f(Q_1, Q_2) \quad (18.19)$$

where π is the firm's profit, Q_1 is its output level of the first good, and Q_2 is its output level of the second good.

To find the value of each of the independent variables (Q_1 and Q_2 in this case) that maximizes the dependent variable (π in this case), we need to know the marginal effect of each independent variable on the dependent variable, *holding constant the effect of all other independent variables*. For example, in this case we need to know the marginal effect of Q_1 on π when Q_2 is held constant, and we need to know the marginal effect of Q_2 on π when Q_1 is held constant. To get this information, we obtain the partial derivative of π with respect to Q_1 and the partial derivative of π with respect to Q_2 .

To obtain the partial derivative of π with respect to Q_1 , denoted $\partial\pi/\partial Q_1$, we apply the rules for finding a derivative (on pages 738–44) to equation (18.19),

4. Two points should be noted.

(1) For profit to be maximized, $d^2\pi/dQ^2$ must be negative. (2) The analysis in this section (as well as in earlier sections) results in the determination of a *local* maximum. Sometimes a local maximum is not a global maximum. For example, under some circumstances, the profit-maximizing (or loss-minimizing) output is zero.

but we treat Q_2 as a constant. Similarly, to obtain the partial derivative of π with respect to Q_2 , denoted $\partial\pi/\partial Q_2$, we apply these rules to equation (18.19), but we treat Q_1 as a constant.

EXAMPLE Suppose the relationship between the Merrimack Company's profit (in thousands of dollars) and its output level of each good is

$$\pi = -20 + 113.75Q_1 + 80Q_2 - 10Q_1^2 - 10Q_2^2 - 5Q_1Q_2 \quad (18.20)$$

To find the partial derivative of π with respect to Q_1 , we treat Q_2 as a constant and find that

$$\frac{\partial\pi}{\partial Q_1} = 113.75 - 20Q_1 - 5Q_2$$

To find the partial derivative of π with respect to Q_2 , we treat Q_1 as a constant and find that

$$\frac{\partial\pi}{\partial Q_2} = 80 - 20Q_2 - 5Q_1$$

Once we have obtained the partial derivatives, it is relatively simple to determine the values of the independent variables that maximize the dependent variable. All we have to do is *set all the partial derivatives equal to zero*. In the case of the Merrimack Company

$$\frac{\partial\pi}{\partial Q_1} = 113.75 - 20Q_1 - 5Q_2 = 0 \quad (18.21)$$

$$\frac{\partial\pi}{\partial Q_2} = 80 - 20Q_2 - 5Q_1 = 0 \quad (18.22)$$

Equations (18.21) and (18.22) are two equations in two unknowns. Solving them simultaneously, we find that profit is maximized when $Q_1 = 5.0$ and $Q_2 = 2.75$. In other words, to maximize profit, the firm should produce 5.0 units of the first good and 2.75 units of the second good per period of time. If it does this, its profit will equal \$374,375 thousand per period of time.⁵

To see why all the partial derivatives should be set equal to zero, consider Figure 18.12, which shows the relationship in equation (18.20) among π , Q_1 , and Q_2 in the range where π is close to its maximum value. As you can see, this relationship is represented by a three-dimensional surface. The maximum value of π is at point M , where this surface is level. A plane tangent to this surface at point M is parallel to the Q_1Q_2 plane; in other words, its slope with respect to either Q_1 or Q_2 must be zero. Because the partial derivatives in equations (18.21) and (18.22) equal these slopes, they too must equal zero at the maximum point M .⁶

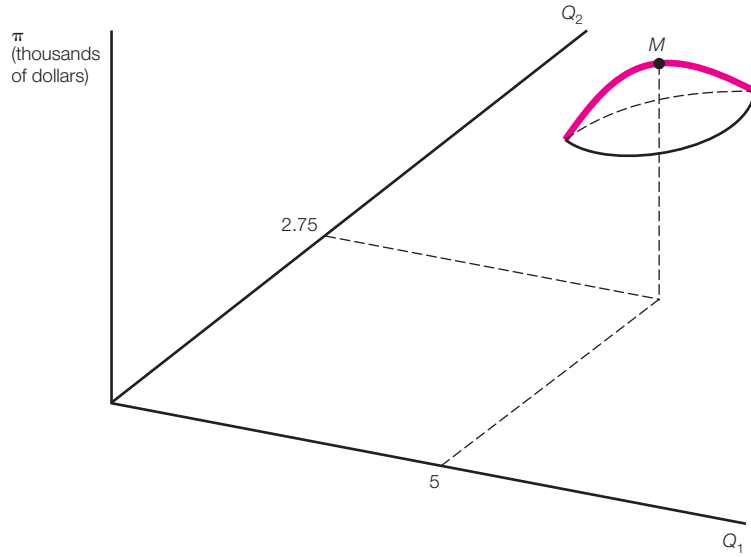
5. Inserting 5.0 for Q_1 and 2.75 for Q_2 in equation (18.20), we find that $\pi = -20 + 113.75(5) + 80(2.75) - 10(5)^2 - 10(2.75)^2 - 5(5)(2.75) = 374.375$

6. The second-order conditions for distinguishing maxima from minima can be found in any calculus book. For present purposes, a discussion of these conditions is not essential.

FIGURE 18.12

Relationship among π , Q_1 , and Q_2

At M , the point where π is a maximum, the surface representing this relationship is flat; its slope with regard to either Q_1 or Q_2 is zero.

**CONSTRAINED OPTIMIZATION**

Managers of firms and other organizations generally face constraints that limit their options. A production manager may want to minimize his or her firm's costs but may not be permitted to produce less than is required to meet the firm's contracts with its customers. The managers of a firm may want to maximize profits; but in the short run, they may be unable to change its product or augment its plant and equipment.

Constrained optimization problems of this sort can be solved in a number of ways. In relatively simple cases in which there is only one constraint, we can use this constraint to express a decision variable—that is, one of the variables the decision maker can choose—as a function of the other decision variables. Then we can apply the techniques for unconstrained optimization described in the previous sections. In effect, we convert the problem to one of unconstrained maximization or minimization.

To illustrate, suppose the Kloster Company produces two products, and its total cost equals

$$TC = 4Q_1^2 + 5Q_2^2 - Q_1Q_2 \quad (18.23)$$

where Q_1 equals its output per hour of the first product and Q_2 equals its output per hour of the second product. Because of commitments to customers, the number produced of both products combined cannot be less than 30 per hour. Kloster's president wants to know what output levels of the two products minimize the firm's costs, given that the output of the first product plus the output of the second product equals 30 per hour.

This constrained optimization problem can be expressed as follows

$$\begin{array}{ll} \text{Minimize} & TC = 4Q_1^2 + 5Q_2^2 - Q_1Q_2 \\ \text{subject to} & Q_1 + Q_2 = 30 \end{array}$$

Of course the constraint is that $(Q_1 + Q_2)$ must equal 30. Solving this constraint for Q_1 , we have

$$Q_1 = 30 - Q_2$$

Substituting $(30 - Q_2)$ for Q_1 in equation (18.23), it follows that

$$\begin{aligned} TC &= 4(30 - Q_2)^2 + 5Q_2^2 - (30 - Q_2)Q_2 \\ &= 4(900 - 60Q_2 + Q_2^2) + 5Q_2^2 - 30Q_2 + Q_2^2 \\ TC &= 3,600 - 270Q_2 + 10Q_2^2 \end{aligned} \quad (18.24)$$

The methods of unconstrained optimization just described can be used to find the value of Q_2 that minimizes TC . As indicated in earlier sections, we must obtain the derivative of TC with respect to Q_2 and set it equal to zero

$$\begin{aligned} \frac{dTC}{dQ_2} &= -270 + 20Q_2 = 0 \\ 20Q_2 &= 270 \\ Q_2 &= 13.5 \end{aligned}$$

To be sure this is a minimum, not a maximum, we obtain the second derivative, which is

$$\frac{d^2TC}{dQ_2^2} = 20$$

Because this is positive, we have found a minimum.

To find the value of Q_1 that minimizes total cost, recall that the constraint requires that

$$Q_1 + Q_2 = 30$$

which means that

$$Q_1 = 30 - Q_2$$

We know that the optimal value of Q_2 is 13.5, so the optimal value of Q_1 must be

$$Q_1 = 30 - 13.5 = 16.5$$

Summing up, if the Kloster Company wants to minimize total cost subject to the constraint that the sum of the output levels of its two products remains 30, it should produce 16.5 units of the first product and 13.5 units of the second product per hour.⁷ In other words, it should produce 33 units of the first product and 27 units of the second product every two hours.

LAGRANGIAN MULTIPLIERS

If the technique described in the previous section is not feasible because the constraints are too numerous or complex, the method of Lagrangian multipliers can be used. This method of solving constrained optimization problems involves the construction of an equation—the so-called Lagrangian function—that combines the function to be minimized or maximized and the constraints. This equation is constructed so that two things are true:

1. When this equation is maximized (or minimized), the original function we want to maximize (or minimize) is in fact maximized (or minimized).
2. All the constraints are satisfied.

To illustrate how we create a Lagrangian function, reconsider the problem faced by the Kloster Company. As indicated in the previous section, this firm wants to minimize $TC = 4Q_1^2 + 5Q_2^2 - Q_1Q_2$, subject to the constraint that $Q_1 + Q_2 = 30$. The first step in constructing the Lagrangian function for this firm's problem is to restate the constraint so that an expression is formed that is equal to zero

$$30 - Q_1 - Q_2 = 0 \quad (18.25)$$

If we multiply this form of the constraint by an unknown factor, designated λ (*lambda*), and add the result to the function we want to minimize (in equation (18.23)), we get the Lagrangian function, which is

$$L_{TC} = 4Q_1^2 + 5Q_2^2 - Q_1Q_2 + \lambda(30 - Q_1 - Q_2) \quad (18.26)$$

For reasons specified in the next paragraph, we can be sure that if we find the unconstrained maximum (or minimum) of the Lagrangian function, the solution will be exactly the same as the solution of the original constrained maximization (or minimization) problem. In other words, to solve the constrained optimization problem, all we have to do is optimize the Lagrangian function. For example, in the case of the Kloster Company, we must find the values of Q_1 , Q_2 , and λ that minimize L_{TC} in equation (18.26). To do this, we must find the partial derivative of L_{TC} with respect to each of the three variables Q_1 , Q_2 , and λ

7. Substituting 16.5 for Q_1 and 13.5 for Q_2 in equation (18.23), we see that the firm's total cost will equal

$$\begin{aligned} TC &= 4(16.5)^2 + 5(13.5)^2 - (16.5)(13.5) \\ &= 4(272.25) + 5(182.25) - 222.75 \\ &= 1089 + 911.25 - 222.75 \\ &= 1,777.5 \text{ or } \$1,777.50 \end{aligned}$$

$$\begin{aligned}\frac{\partial L_{TC}}{\partial Q_1} &= 8Q_1 - Q_2 - \lambda \\ \frac{\partial L_{TC}}{\partial Q_2} &= -Q_1 + 10Q_2 - \lambda \\ \frac{\partial L_{TC}}{\partial \lambda} &= -Q_1 - Q_2 + 30\end{aligned}$$

As indicated earlier, we must set all three of these partial derivatives equal to zero to minimize L_{TC}

$$8Q_1 - Q_2 - \lambda = 0 \quad (18.27)$$

$$-Q_1 + 10Q_2 - \lambda = 0 \quad (18.28)$$

$$-Q_1 - Q_2 + 30 = 0 \quad (18.29)$$

It is important to note that the partial derivative of the Lagrangian function with regard to λ (that is, $\partial L_{TC}/\partial \lambda$), when it is set equal to zero (in equation (18.29)), is the constraint in our original optimization problem (recall equation (18.25)). This, of course, is always true because of the way the Lagrangian function is constructed. So if this derivative is zero, we can be sure this original constraint is satisfied. And if this constraint is satisfied, the last term on the right of the Lagrangian function is zero; so the Lagrangian function boils down to the original function that we wanted to maximize (or minimize). Consequently, by maximizing (or minimizing) the Lagrangian function, we solve the original constrained optimization problem.

Returning to the Kloster Company, equations (18.27), (18.28), and (18.29) are three simultaneous equations with three unknowns— Q_1 , Q_2 , and λ . If we solve this system of equations for Q_1 and Q_2 , we get the optimal values of Q_1 and Q_2 . Subtracting equation (18.28) from equation (18.27), we find that

$$9Q_1 - 11Q_2 = 0 \quad (18.30)$$

Multiplying equation (18.29) by 9 and adding the result to equation (18.30), we can solve for Q_2

$$\begin{array}{r} -9Q_1 - 9Q_2 + 270 = 0 \\ \quad 9Q_1 - 11Q_2 = 0 \\ \hline -20Q_2 + 270 = 0 \\ Q_2 = 270/20 = 13.5 \end{array}$$

Therefore, the optimal value of Q_2 is 13.5. Substituting 13.5 for Q_2 in equation (18.29), we find the optimal value of Q_1 is 16.5.

The answer we get is precisely the same as in the previous section: The optimal value of Q_1 is 16.5, and the optimal value of Q_2 is 13.5. In other words, the

managers at Kloster Company should produce 16.5 units of the first product and 13.5 units of the second product per hour. But the method of Lagrangian multipliers described in this section is more powerful than that described in the previous section for at least two reasons: (1) It can handle more than a single constraint, and (2) the value of λ provides interesting and useful information to the decision maker.

Specifically λ , called the *Lagrangian multiplier*, measures the change in the variable to be maximized or minimized (TC in this case) if the constraint is relaxed by one unit. For example, if managers at the Kloster Company want to minimize total cost subject to the constraint that the total output of both products is 31 rather than 30, the value of λ indicates by how much the minimum value of TC will increase. What is the value of λ ? According to equation (18.27)

$$8Q_1 - Q_2 - \lambda = 0$$

Because $Q_1 = 16.5$ and $Q_2 = 13.5$

$$\lambda = 8(16.5) - 13.5 = 118.5$$

Consequently, if the constraint is relaxed so that total output is 31 rather than 30, the total cost will go up by \$118.50.

For many managerial decisions, information of this sort is of great value. Suppose a customer offers the Kloster Company \$115 for one of its products, but to make this product, Kloster would have to stretch its total output to 31 per hour. On the basis of the findings of the previous paragraph, Kloster managers would be foolish to accept this offer because this extra product will raise costs by \$118.50, \$3.50 more than the amount the customer offers to pay for it.

COMPARING INCREMENTAL COSTS WITH INCREMENTAL REVENUES

Before concluding this chapter, we must point out that many business decisions require comparing incremental costs with incremental revenues. Typically a manager must choose among courses of action, and the relevant decision factors are the differences in costs and revenues. For example, if the managers of a machinery company are considering whether to add a new product line, they should compare the incremental cost of adding the new product line (the extra cost resulting from its addition) with the incremental revenue (the extra revenue resulting from its addition). If the incremental revenue exceeds the incremental cost, the new product line will add to the firm's profits.

Note that *incremental* cost is not the same as *marginal* cost. Whereas marginal cost is the extra cost from a very small (one-unit) increase in output, **incremental cost** is the extra cost from an output increase that may be substantial. Similarly, **incremental revenue**, unlike marginal revenue, is the extra revenue from an output increase that may be substantial. For example, suppose you want to see

Incremental cost The extra cost from an output increase that may be substantial.

Incremental revenue The extra revenue from an output increase that may be substantial.

whether a firm's profits will increase if it doubles its output. If the incremental cost of such an output increase is \$5 million and the incremental revenue is \$6 million, the firm will increase its profits by \$1 million if it doubles its output. Marginal cost and marginal revenue cannot tell you this because they refer to only a very small increase in output, not to a doubling of it.

Although it may seem easy to compare incremental costs with incremental revenues, in fact there are many pitfalls. One of the most common errors is the failure to recognize the irrelevance of sunk costs. Costs incurred in the past are often irrelevant in making today's decisions. Suppose you are going to make a trip and you want to determine whether it will be cheaper to drive your car or to travel by plane. What costs should be included if you drive your car? Because the only incremental costs incurred are the gas and oil (and a certain amount of wear and tear on tires, engine, and so on), these are the only costs to be included. Costs incurred in the past, such as the original price of the car, and costs that are the same regardless of whether you make the trip by car or plane, such as your auto insurance, should not be included. On the other hand, if you are thinking about buying a car to make this and many other trips, these costs should be included.

To illustrate the proper reasoning, consider an airline that has deliberately run extra flights that return only a little more than their out-of-pocket costs. Assume this airline faces the decision of whether to run an extra flight between city *A* and city *B*. The fully allocated costs—the out-of-pocket costs plus a certain percentage of overhead, depreciation, insurance, and other such costs—are \$5,500 for the flight. The out-of-pocket costs—the actual sum this airline has to disburse to run the flight—are \$3,000, and the expected revenue from the flight is \$4,100. In such a case, this airline will run the flight, which is the correct decision because the flight will add \$1,100 to profit. The incremental revenue from the flight is \$4,100, and the incremental cost is \$3,000. Overhead, depreciation, and insurance are the same whether the flight is run or not. Therefore, fully allocated costs are misleading here; the relevant costs are out-of-pocket, not fully allocated, ones.

Errors of other kinds can also mar managerial estimates of incremental costs. For example, a firm may refuse to produce and sell some items because it is already working near capacity, and the incremental cost of producing them is judged to be high. In fact, however, the incremental cost may not be so high because managers may be able to produce these items during the slack season (when there is plenty of excess capacity), and the potential customers may be willing to accept delivery then.

Also, incremental revenue is frequently misjudged. Consider managers who are pondering the introduction of a new product. The firm's managers may estimate the incremental revenue from the new product without taking proper account of the effects of the new product's sales on the sales of existing products. They may think the new product will not cut into the sales of existing products; however, it may in fact do so, with the result that their estimate of incremental revenue is too high.

SUMMARY

1. Functional relationships can be represented by tables, graphs, or equations. The marginal value of a dependent variable is defined as the change in this variable associated with a 1-unit change in a particular independent variable. The dependent variable achieves a maximum when its marginal value shifts from positive to negative.

2. The derivative of Y with respect to X , denoted dY/dX , is the limit of the ratio $\Delta Y/\Delta X$ as ΔX approaches zero. Geometrically, it is the slope of the curve showing Y (on the vertical axis) as a function of X (on the horizontal axis). We have provided rules that enable us to find the value of this derivative.

3. To find the value of X that maximizes or minimizes Y , we determine the value of X where dY/dX equals zero. To tell whether this is a maximum or a minimum, we find the second derivative of Y with respect to X , denoted d^2Y/dX^2 , which is the derivative of dY/dX . If this second derivative is negative, we have found a maximum; if it is positive, we have found a minimum.

4. A dependent variable often depends on a number of independent variables, not just one. To find the value of each of the independent variables that maximizes the dependent variable, we determine the partial derivative of Y with respect to each of the independent variables, denoted $\partial Y/\partial X$, and set it equal to zero. To obtain the partial derivative of Y with respect to X , we apply the ordinary rules for finding a derivative; however, all independent variables other than X are treated as constants.

5. Managers of firms and other organizations generally face constraints that limit the options available to them. In relatively simple cases in which there is only one constraint, we can use this constraint to express one of the decision variables as a function of the other decision variables, and we can apply the techniques for unconstrained optimization.

6. In more complex cases, constrained optimization problems can be solved by the method of Lagrangian multipliers. The Lagrangian function combines the function to be maximized or minimized and the constraints. To solve the constrained optimization problem, we optimize the Lagrangian function.

7. Many business decisions can and should be made by comparing incremental costs with incremental revenues. Typically, a manager must choose between two (or more) courses of action, and what is relevant is the difference between the costs of the two courses of action, as well as the difference between their revenues.

PROBLEMS

1. One very important question facing hospitals is this: How big must a hospital be (in terms of patient-days of care) to minimize the cost per patient-day?

According to one well-known study, the total cost (in dollars) of operating a hospital (of a particular type) can be approximated by

$$C = 4,700,000 + 0.00013X^2$$

where X is the number of patient-days.

- a. Derive a formula for the relationship between cost per patient-day and the number of patient-days.
 - b. On the basis of the results of this study, how big must a hospital be (in terms of patient-days) to minimize the cost per patient-day?
 - c. Show that your result minimizes, rather than maximizes, the cost per patient-day.
2. The Trumbull Company has developed a new product. Trumbull's chairperson estimates that the new product will increase the firm's revenues by \$5 million per year, and that it will result in extra out-of-pocket costs of \$4 million per year, the fully allocated costs (including a percentage of overhead, depreciation, and insurance) being \$5.5 million.
- a. Trumbull's chairperson feels that it would not be profitable to introduce this new product. Is the chairperson right? Why or why not?
 - b. Trumbull's vice president for research argues that since the development of this product has already cost about \$10 million, the firm has little choice but to introduce it. Is the vice president right? Why or why not?
3. For the Martin Corporation, the relationship between profit and output is the following:

Output (number of units per day)	Profit (thousands of dollars per day)
0	-10
1	-8
2	-5
3	0
4	2
5	7
6	12
7	21
8	22
9	23
10	20

- a. What is the marginal profit when output is between 5 and 6 units per day? When output is between 9 and 10 units per day?
- b. At what output is average profit a maximum?

- c. Should the Martin Corporation produce the output where average profit is a maximum? Why or why not?
4. Determine the first derivative of each of the following functions:
 - a. $Y = 3 + 10X + 5X^2$
 - b. $Y = 2X(4 + X^3)$
 - c. $Y = 3X/(4 + X^3)$
 - d. $Y = 4X/(X - 3)$
5. The total cost function at the Duemer Company is $TC = 100 + 4Q + 8Q^2$ where TC is total costs, and Q is the output.
 - a. What is marginal cost when output is 10?
 - b. What is marginal cost when output is 12?
 - c. What is marginal cost when output is 20?
6. The Bartholomew Company's profit is related in the following way to its output: $\pi = -40 + 20Q - 3Q^2$, where π is total profit and Q is output.
 - a. If the firm's output equals 8, what is its marginal profit?
 - b. Derive an equation relating the firm's marginal profit to its output.
 - c. What output maximizes the firm's profit?
7. Determine the second derivative of the following functions:
 - a. $Y = 4 + 9X + 3X^2$
 - b. $Y = 4X(3 + X^2)$
 - c. $Y = 4X(2 + X^3)$
 - d. $Y = (4/X) + 3$
8. The Mineola Corporation hires a consultant to estimate the relationship between its profit and its output. The consultant reports that the relationship is

$$\pi = -10 - 6Q + 5.5Q^2 - 2Q^3 + 0.25Q^4$$

- a. The consultant says that the firm should set Q equal to 1 to maximize profit. Is it true that $d\pi/dQ = 0$ when $Q = 1$? Is π at a maximum when $Q = 1$?
- b. Mineola's executive vice president says that the firm's profit is a maximum when $Q = 2$. Is this true?
- c. If you were the chief executive officer of the Mineola Corporation, would you accept the consultant's estimate of the relationship between profit and output as correct?
9. Find the partial derivative of Y with respect to X in each of the following cases:
 - a. $Y = 10 + 3Z + 2X$
 - b. $Y = 18Z^2 + 4X^3$
 - c. $Y = Z^{0.2}X^{0.8}$
 - d. $Y = 3Z/(4 + X)$
10. The Stock Corporation makes two products, paper and cardboard. The relationship between π , the firm's annual profit (in thousands of dollars), and its output of each good is

$$\pi = -50 + 40Q_1 + 30Q_2 - 5Q_1^2 - 4Q_2^2 - 3Q_1Q_2$$

where Q_1 is the firm's annual output of paper (in tons), and Q_2 is the firm's annual output of cardboard (in tons).

- a. Find the output of each good that the Stock Corporation should produce if it wants to maximize profit.
 - b. If the community in which the firm is located imposes a tax of \$5,000 per year on the firm, will this alter the answer to Part a? If so, how will the answer change?
11. The Miller Company uses skilled and unskilled labor to do a particular construction project. The cost of doing the project depends on the number of hours of skilled labor and the number of hours of unskilled labor that are used, the relationship being

$$C = 4 - 3X_1 - 4X_2 + 2X_1^2 + 3X_2^2 + X_1X_2$$

where C is cost (in thousands of dollars), X_1 is the number of hours (in thousands) of skilled labor, and X_2 is the number of hours (in thousands) of unskilled labor.

- a. Find the number of hours of skilled labor and the number of hours of unskilled labor that will minimize the cost of doing the project.
 - b. If the Miller Company has to purchase a license costing \$2,000 to do this project (and if the cost of this license is not included in C), will this alter the answer to Part a? If so, how will the answer change?
12. Ilona Stafford manages a small firm that produces wool rugs and cotton rugs. Her total cost per day (in dollars) equals

$$C = 7X_1^2 + 9X_2^2 - 1.5X_1X_2$$

where X_1 equals the number of cotton rugs produced per day, and X_2 equals the number of wool rugs produced per day. Because of commitments to retail stores that sell her rugs to consumers, she must produce ten rugs per day, but any mix of wool and cotton rugs is acceptable.

- a. If she wants to minimize her costs (without violating her commitment to the retail stores), how many cotton rugs and wool rugs should she produce per day? (Do *not* use the method of Lagrangian multipliers.)
 - b. Does it seem reasonable that she would want to minimize cost in a situation of this sort? Why or why not?
 - c. Can she produce fractional numbers of rugs per day?
13. a. Use the method of Lagrangian multipliers to solve Problem 12.
- b. Do you get the same answer as you do without using this method?
 - c. What does λ equal? What does this mean?

TECHNOLOGICAL CHANGE AND INDUSTRIAL INNOVATION

Companies in the United States have traditionally been at the forefront in using new technologies. Their long-term profitability and market successes are often based on new products or processes. Most of these advances in products or processes are the result of an active commitment to devoting resources to research and development. After all, increases in productivity do not occur haphazardly but are the result of active management of the research and development process. This management not only includes the initial development of a product or process but also the implementation issues involved in getting new products to market.

In this appendix, we discuss various models and techniques for measuring productivity and examine several research and developmental models. These models have proven useful in both helping managers oversee research and development programs and bringing new products to market.

TECHNOLOGICAL CHANGE

Technological change—the advance of technology—often takes the form of new methods of producing existing products and new techniques of organization, marketing, and management. Technological change results in a change in the production function. If the production function were readily observable, a

LEARNING OBJECTIVES

- Technological Change
- Labor Productivity
- Total Factor Productivity
- Using Total Factor Productivity to Track Factory Performance
- Research and Development: A Learning Process
- Parallel Development Efforts
- What Makes for Success?
- Project Selection
- Innovation
- Time-Cost Trade-offs
- The Learning Curve
- Applications of the Learning Curve
- Henry Ford's Model T and Douglas Aircraft's DC-9
- Diffusion Models
- Forecasting the Rate of Diffusion of Numerically Controlled Machine Tools
- Summary
- Problems

Technological change Results in a change of a firm's production function over time. This can entail existing products produced more efficiently or can result in the availability of new products.

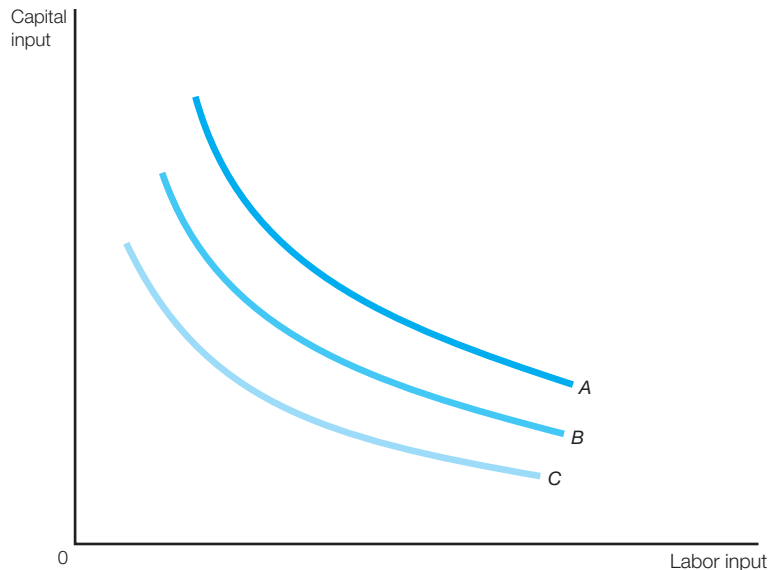
comparison of the production function at two different times would provide the manager with a simple measure of the effect of technological change during the intervening period. If there were only two inputs, labor and capital, and constant returns to scale, the characteristics of the production function at a given date could be captured fully by a single isoquant.¹ One could simply look at the changing position of this isoquant to see the effects of technological change. If this isoquant shifted from position *A* to position *B* in Figure A.1 during a certain period of time, technological change had less impact during this period than it would have had if the curve had shifted to position *C*.

Technological change may also result in the availability of new products. DVD players, for example, did not exist several decades ago; now they are commonplace. iPhones and tablets did not exist a decade ago; today, many managers rely on them. Nylon was first brought to market in the 1930s; today it is hard to imagine what life would be like without it. In many cases, the availability of new products can be regarded as a change in the production function, since they are merely more efficient ways of meeting old wants if these wants are defined with proper breadth. This is particularly true in the case of new goods used by firms, which may result in little or no change in the final product shipped to consumers.

FIGURE A.1

Change over a Period of Time in the Position of an Isoquant

If the isoquant shifted from position *A* to position *B*, technological change had less impact than if it had shifted to position *C*.



1. Recall that, if there are constant returns to scale, an x percent increase in all inputs results in an x percent increase in output. Hence, if there are constant returns to scale, there is at a given time a unique relationship between capital input per unit of output and labor input per unit of output. This relationship holds for any output and completely summarizes the efficient input combinations.

In other cases, however, the availability of new products cannot realistically be viewed as a change in the production function, since the new products represent an important difference in kind.

LABOR PRODUCTIVITY

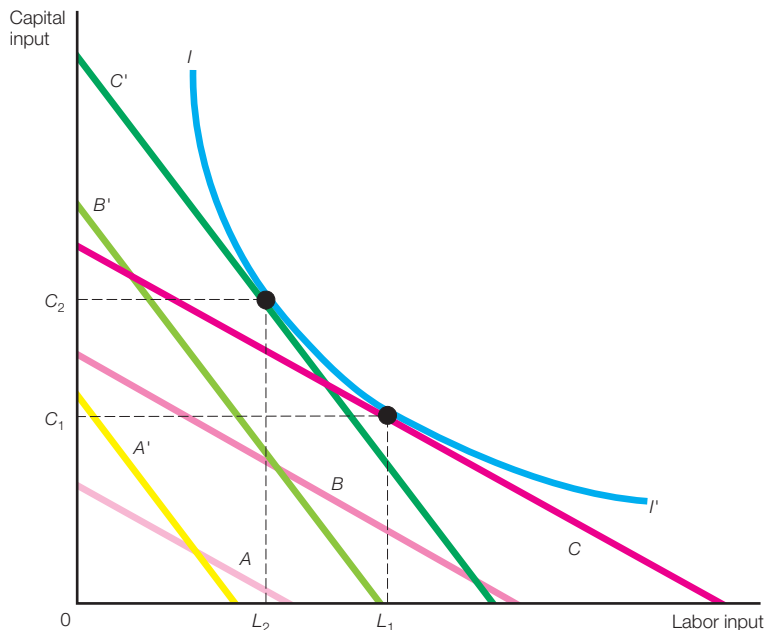
Managers have long been interested in productivity—the ratio of output to input. The oldest and most commonly studied productivity measure is labor productivity, output per hour of labor. One determinant of the rate of growth of labor productivity is the rate of technological change: A high rate of technological change is likely to result, all other things being equal, in a high rate of growth of labor productivity. However, the rate of technological change is not the sole determinant of the rate of growth of labor productivity; as a consequence, although labor productivity is often used to measure the rate of technological change, it is in fact an incomplete measure.

Figure A.2 shows how changes in labor productivity can produce false signals concerning the rate of technological change. Suppose the relevant isoquant

FIGURE A.2

Productivity Increase without Technological Change

Because labor becomes more expensive relative to capital, labor productivity increases.



is II' and the input prices at the beginning of the period are such that the isocost curves are A, B, C , and so on. The least-cost combination of inputs is L_1 of labor and C_1 of capital. Now suppose that input prices change and labor becomes more expensive relative to capital; as a result, the isocost curves shift to A', B', C' , and so on. Under these new circumstances, the least-cost combination of inputs to produce the same output is L_2 of labor and C_2 of capital. Since output remains constant and labor input decreases, labor productivity increases as a result of the change in input prices. But this productivity increase is not an indication of technological change, there being no change at all in the production function.

TOTAL FACTOR PRODUCTIVITY

Total factor productivity relates changes in the firm's outputs to changes in all the firm's inputs.

A better measure of the rate of technological change is **total factor productivity**, which relates changes in output to changes in both labor and capital inputs, not changes in labor inputs alone. Assume that the production function is of the simple form

$$Q = \alpha(bL + cK) \quad (\text{A.1})$$

where Q is the quantity of output, L is the quantity of labor, K is the quantity of capital, and b and c are constants. Dividing both sides of equation (A.1) by $(bL + cK)$

$$\frac{Q}{bL + cK} = \alpha \quad (\text{A.2})$$

which is total factor productivity. In this simple case, changes in total factor productivity measure changes in efficiency.

If a firm uses more than two inputs, total factor productivity equals

$$\frac{Q}{a_1 I_1 + a_2 I_2 + \dots + a_n I_n} \quad (\text{A.3})$$

where I_1 is the amount of the first input used, I_2 is the amount of the second input used, \dots , and I_n is the amount of the n th input used. In calculating total factor productivity, firms often let a_1 equal the price of the first input, a_2 equal the price of the second input, \dots , and a_n equal the price of the n th input in some base period, as we shall see next. The principal advantage of total factor productivity over labor productivity is its inclusion of more types of inputs, not labor alone. It otherwise shares many of the limitations of labor productivity. Note that $\sum_i P_i X_i$, where P_i is unit input cost and X_i is amount of input used, is just the firm's total cost (TC). Then Q/TC is just the inverse of the firm's average total cost (ATC). To maximize productivity, minimize ATC. But firms want to maximize profits. That rarely occurs when firms minimize ATC.

Firms calculate total factor productivity to measure changes over time in the efficiency of their operations. It is important for a firm's managers to be aware of the extent to which productivity has increased in response to new techniques and other factors.

To calculate the changes in total factor productivity for a firm or plant over a period of time, managers must obtain data concerning the quantities of output and inputs utilized in each period. For example, suppose that the Landau Company uses three inputs—labor, energy, and raw materials. In 2012, it uses 10,000 hours of labor, 100,000 kilowatt-hours of energy, and 5,000 pounds of materials to produce 400,000 pounds of output. In 2013, it uses 12,000 hours of labor, 150,000 kilowatt-hours of energy, and 6,000 pounds of materials to produce 700,000 pounds of output. What is total factor productivity in each year?

As a first step toward answering this question, we must get data concerning the price of each input in some base period, say, 2012. Suppose that the price of labor is \$8 per hour, the price of a kilowatt-hour of energy is \$0.02, and the price of a pound of materials is \$3. Then, inserting these figures into the expression in (A.3), we find that total factor productivity in 2012 is

$$\frac{400,000}{8(10,000) + 0.02(100,000) + 3(5,000)} = 4.12$$

and total factor productivity in 2013 is

$$\frac{700,000}{8(12,000) + 0.02(150,000) + 3(6,000)} = 5.98$$

Therefore, from 2012 to 2013, total factor productivity increased by 45%—from 4.12 to 5.98.

Note that the base-year input prices are used for all years, not just the base year. For example, the 2012 input prices would be used for all years, not just 2012, in the case of the Landau Company. In this way, we hold constant input prices and do not let changes in them over time affect our results.²

USING TOTAL FACTOR PRODUCTIVITY TO TRACK FACTORY PERFORMANCE

To illustrate how changes in total factor productivity can be used to track factory performance, consider a manufacturing plant studied by Harvard's Robert Hayes, Steven Wheelwright, and Kim Clark.³ Figure A.3 shows the behavior of total factor productivity in this plant during a 10-year period. As you can see, total factor productivity increased at a healthy pace up to 1976. This was the period during which the plant was started up. Because it takes time for a factory to operate properly, one would expect that total factor productivity would increase substantially in this startup phase.

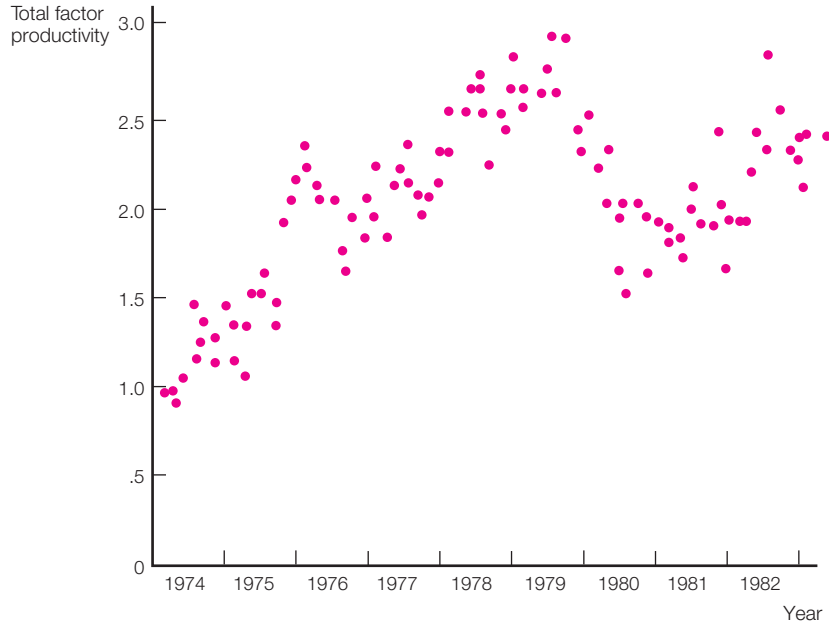
2. Of course, this does not mean that the value of total factor productivity is not affected by the base-year prices. For example, if the price of labor in the base period were \$10 (rather than \$8) per hour, our results would be different. But changes *over time* in input prices are not allowed to influence our results.

3. R. Hayes, S. Wheelwright, and K. Clark, *Dynamic Manufacturing* (New York: Free Press, 1988).

FIGURE A.3

Total Factor Productivity, Actual Manufacturing Plant

Total factor productivity increased up to 1976, but in 1982 was only slightly higher than in 1976.



Source: R. Hayes, S. Wheelwright, and K. Clark, *Dynamic Manufacturing*. (New York: Free Press, 1988).

From 1977 to 1983, there was no evidence of any strong, persistent increase in total factor productivity. Instead, there was an increase in 1977 to 1979, a fall in 1979 to 1980, and an increase in 1981 to 1982. In 1982, total factor productivity was only somewhat higher than in 1976. The data in Figure A.3 indicate that this factory experienced little in the way of technological change from 1977 to 1983. Hayes, Wheelwright, and Clark report that these data triggered an investigation of the causes for this poor performance, which indicated that it was due in considerable part to the way the factory managed equipment introductions.⁴ Obviously, this information was of use to the firm's top managers.

In passing, note the fact that total factor productivity fell during 1979 to 1980 does not mean that there was negative technological change then. If the factory's sales decreased during this period, perhaps because of cyclical factors, this could cause such a decline in total factor productivity. Also, it sometimes takes time for new equipment to reach its full efficiency. When equipment is first used, productivity may fall temporarily because of "teething" problems.

4. Ibid.

RESEARCH AND DEVELOPMENT: A LEARNING PROCESS

Particularly in science-based industries like electronics and chemicals, a firm's success depends on the extent and nature of the research and development that it carries out. Research and development encompasses work of many kinds. Basic research is aimed purely at the creation of new knowledge, applied research is expected to have a practical payoff, and development is aimed at the reduction of research findings to practice. Inventions can occur in either the research phase or the development phase of organized research and development activity.

Chance plays a crucial role in research and development, and a long string of failures frequently occurs before any kind of success is achieved. A research or development project can be regarded as a process of uncertainty reduction, or learning. Suppose, for example, a manager who is trying to fabricate a part can use one of two alloys and it is impossible to use standard sources to determine their characteristics. Suppose that strength is of paramount importance and the manager's estimates of the strengths of the alloys, alloy *X* and alloy *Y*, are represented by the probability distributions in part A of Table A.1. If the manager was forced to make a choice immediately, she would probably choose alloy *Y*, since she believes there is better than a 50–50 chance that alloy *Y* will turn out to be stronger than alloy *X*.

However, there is a good chance that this decision might turn out to be wrong, with the consequence that the part would be weaker than if alloy *X* had been used. Therefore, the manager may decide to perform a test prior to making the selection. On the basis of the test results, the manager formulates new estimates, represented by the probability distributions in part B of Table A.1. These probability

TABLE A.1

Subjective Probability Distribution of Strength of Alloys *X* and *Y*

Extent of strength	Probabilities			
	A. Before test		B. After test	
	Alloy <i>X</i>	Alloy <i>Y</i>	Alloy <i>X</i>	Alloy <i>Y</i>
Exceptionally high	0.20	0.30	0.10	0.10
Very high	0.40	0.50	0.20	0.80
High	0.20	0.10	0.60	0.10
Medium	0.10	0.05	0.10	0.00
Low	0.10	0.05	0.00	0.00
Total	1.00	1.00	1.00	1.00

distributions show less dispersion than the distributions in part A; in other words, the manager believes she is able to pinpoint more closely the strength of each alloy in part B than in part A. Because of the tests, the manager feels more certainty concerning which alloy will prove stronger.

PARALLEL DEVELOPMENT EFFORTS

Research and development is more risky than most other economic activities. Many development projects use parallel efforts to help cope with the uncertainty. For example, in the development of the atomic bomb, there were several methods of making fissionable materials and no consensus among scientists as to which of these alternatives was the most promising. To make sure that the best one was not discarded, all methods were pursued in parallel. The wisdom of this decision was borne out by the fact that the first method to produce appreciable quantities of fissionable material was one considered relatively unpromising early in the development program's history.

How can a firm's managers tell whether it is optimal to run parallel research and development efforts? What factors determine the optimal number of parallel efforts? Suppose that a firm can select x approaches, spend C dollars on each one over a period of n months, choose the one that looks most promising at the end of the period, and carry it to completion, dropping the others. Suppose that the only relevant criterion is the extent of the development costs, the usefulness of the result and the development time, assumed to be the same regardless of which parallel effort is pursued. For further simplification, suppose that all approaches look equally promising. Under these circumstances, the optimal value of x (the number of parallel research and development efforts) is inversely related to C and directly related to the amount learned in the next n months. As the cost of running each effort increases, the optimal number of parallel efforts decreases. As the prospective amount of learning increases, the optimal number of parallel efforts goes up.

To illustrate why it is sometimes cheaper to run parallel development efforts, consider a case in which each approach has a 50–50 chance of costing \$5 million and a 50–50 chance of costing \$8 million. Since we assume that all approaches are equally promising, these probabilities are the same for all approaches. The expected total cost of development is the sum of the total costs of development if each possible outcome occurs times the probability of the occurrence of this outcome. If a single approach is used, the expected total costs of development are

$$0.5(\$5 \text{ million}) + 0.5(\$8 \text{ million}) = \$6.5 \text{ million} \quad (\text{A.4})$$

since there is a 0.5 probability that total costs with any single approach will be \$5 million and a 0.5 probability that they will be \$8 million.

If two approaches are run in parallel and if the true cost of development using each approach can be determined after C dollars are spent on each approach, the expected total costs of development are

$$0.25(\$8 \text{ million}) + 0.75(\$5 \text{ million}) + C = \$5.75 \text{ million} + C \quad (\text{A.5})$$

If each approach is carried to the point at which C dollars have been spent on it, the cheaper approach is chosen at that point (the other approach is dropped). Why? Because there is a 0.25 probability that total costs with the better of the two approaches will be \$8 million and a 0.75 probability that they will be \$5 million. In addition, there is the certainty that a cost of C will be incurred for the approach that is dropped. (The C dollars spent on the project that is not dropped are included in its total costs, given previously.) The reason why there is a 0.25 chance that total costs with the better of the two approaches is \$8 million is that this will occur only when the total cost of both approaches turns out to be \$8 million—and the probability that this will occur is 0.5 times 0.5, or 0.25. Comparing equation (A.4) with equation (A.5), it is obvious that the expected total cost of development is lower with two parallel approaches than with a single approach if C is less than \$750,000.

More generally, if the probability is P that the development cost will be C_1 and $(1 - P)$ that it will be C_2 (where $C_2 < C_1$), the expected cost if a single approach is used is

$$PC_1 + (1 - P)C_2$$

If two approaches are run in parallel, the expected cost is

$$P^2C_1 + (1 - P^2)C_2 + C$$

which is less than the cost of a single approach if

$$C < (1 - P)(P)(C_1 - C_2) \quad (\text{A.6})$$

Therefore, if the inequality in (A.6) holds, two parallel approaches result in a lower expected cost than a single approach.

WHAT MAKES FOR SUCCESS?

Even companies in the same industry may differ markedly in their ability to make research and development (R and D) pay off commercially. During a four-year period, for instance, three evenly matched chemical companies found the proportion of their R and D expenditures that earned a profit to be 69%, 54%, and 39%, respectively. These differences are too large to be attributed to errors of measurement or definition. What can explain them?⁵

An R and D project's likelihood of economic success is the product of three factors: (1) the probability of technical success, (2) the probability of

5. For references and sources of information for the data presented in this and the next three sections, see E. Mansfield, "How Economists See R and D," *Harvard Business Review*, November–December 1981. Also, see K. Clark and T. Fujimoto, *Product Development Performance* (Boston: Harvard Business School Press, 1991); R. Stobaugh, *Innovation and Competition* (Boston: Harvard Business School Press, 1988); and E. Mansfield, *Innovation, Technology, and the Economy* (Aldershot, U.K.: Elgar, 1995).

commercialization (given technical success), and (3) the probability of economic success (given commercialization). One econometric study shows that all three of these probabilities are directly related to how quickly an R and D project is evaluated for its economic, as opposed to technical, potential. Also, in those companies whose R and D staff members do not work closely or responsively with the marketing staff, the integration of R and D activity with market realities is haphazard, belated, or both. Commercially successful innovation depends on just this sort of integration. Numerous case studies of successful and unsuccessful innovation come to the same conclusion: The closer is the link between marketing and R and D, the greater is the probability of commercialization (given technical completion).

Consider, by way of illustration, the experience of three chemical companies of roughly the same size and level of R and D expenditure that underwent reorganization at roughly the same time. In two of them, the reorganization produced a closer integration of R and D with marketing by improving the channels of communication between them as well as by noticeably increasing marketing's input to R and D decision making. In the third, however, integration decreased; R and D paid even less attention to marketing than it had before the reorganization.

Data on the probability of commercialization (given technical completion) of 330 R and D projects in these companies (projects carried out anywhere from three to seven years before reorganization to five to eight years after it) are highly suggestive. They show an increase of about 20 percentage points for the two companies that more closely linked R and D with marketing and a decrease of about 20 percentage points for the third.

More generally, a substantial portion of a company's R and D efforts may lie fallow because other parts of the company do not make proper use of them. One survey of executive opinion has noted the widely held belief that the economic success rate of R and D projects would increase by half if marketing and production people fully exploited them. If this figure is anywhere close to the truth, the faulty interface between R and D and the other functions has a very serious effect on the productivity of industrial R and D.

PROJECT SELECTION

However well-founded the fears of excessively detailed control, some managerial oversight of R and D is essential. To make effective use of its R and D capacity, a company must spell out its business objectives and communicate them to its scientists and engineers. Research, after all, makes sense only when undertaken in areas relevant to economic goals.

Simply taking on a team of scientists and allowing them to do research in their favorite fields may produce novel results but results that are unlikely to have much immediate commercial value. Most companies, therefore, have found it

worthwhile to make economic evaluations of both project proposals and continuing projects. Without question, these evaluations are useful, since they force managers to make their assumptions explicit. Research suggests that the sooner such evaluations are carried out, the greater a project's chances of ultimate commercial success.

The nature of these evaluations is different for a research lab rather than a development project. As a project moves from the laboratory toward the market, it receives more intensive scrutiny from both the technical and economic angles. In the early research phase, the screening of proposals probably is quick and informal, since costs at this stage are still low and predicting outcomes is very difficult. But, as projects enter the development phase, where costs and predictability are higher, they require a far more detailed process of economic evaluation.

Managerial economists have developed a number of more or less sophisticated models to help solve these problems of evaluation. Some employ relatively straightforward adaptations of capital budgeting techniques. For example, the net present value or internal rate of return (concepts developed in a basic course in finance) of each project may be calculated and compared. The more complicated versions of these models have not found extensive use, for the following

STRATEGY SESSION: Parallel Development Efforts at IBM

The IBM Corporation, which spends billions of dollars per year on research and development, is one of the world's leading high-technology companies. Nonetheless, IBM, like other firms, must face the fact that R and D is a risky activity: It is not able to predict with confidence whether a particular R and D project will be successful. Recognizing this fact, parallel development efforts have played a major role in IBM's history, as indicated by the following quotation from one IBM manager:

Parallel projects are crucial—no doubt of it. When I look back over the last dozen products we've introduced, I find in well over half the instances the big development project that we "bet on" via the system came a cropper somewhere along the way. In every instance—and we've gone back and taken a look and I do mean

every—there were two or three (about five once) other small projects, you know, four-to-six person groups, two people in one instance, who had been working on parallel technology or parallel development efforts. It had been with scrounged time and bodies. But that's a time-honored thing. We wink at it. It pays off. Looking at the projects where the initial bets failed, the subsequently developed project came in ahead of the original schedule in three instances. It's just amazing what a handful of dedicated people can do when they are really turned on. Of course they had an advantage. Since they were so resource-constrained, they had to design a simple product in the first place.^a

^a Bartlett, *Cases in Strategic Management*. (Dryden Press, 1988).

reasons: (1) Many of the models fail to recognize that R and D is essentially a process of buying information, unsuccessful projects can provide valuable information, and as a result, the real task is to facilitate sequential decision making under conditions of uncertainty. (2) Application of the more-sophisticated models is not cheap. (3) Perhaps most important, the models often rest on overly optimistic estimates that are not very reliable—estimates that reflect both the uncertainty of the undertaking and the desire by researchers and others to “sell” projects to top management.

INNOVATION

An **innovation** occurs when an invention is applied for the first time.

An invention, when applied for the first time, is called an **innovation**. The distinction between an invention and an innovation becomes somewhat blurred in cases like DuPont’s nylon, in which the inventor and the innovator are the same firm. In these circumstances, the final stages of development may entail at least a partial commitment to a market test. However, in many cases, the firm that is the inventor is not in a position to—and does not want to—apply its invention, because its business is invention, not production; because it is a supplier, not a user, of the equipment embodying the innovation; or for some other reason. In these cases, the distinction remains relatively clear-cut.

Regardless of whether the break between invention and innovation is clean, innovation is a key stage in the process leading to the full evaluation and utilization of an invention. The innovator—the firm that is first to apply the invention—must be willing to take the risks involved in introducing a new and untried process, good, or service. In many cases, these risks are high. Although R and D can provide a great deal of information regarding the technical characteristics and cost of production of the invention—and market research can provide considerable information regarding the demand for it—many areas of uncertainty can be resolved only by actual production and marketing of the invention. By obtaining needed information regarding the actual performance of the invention, the innovator plays a vital social role.

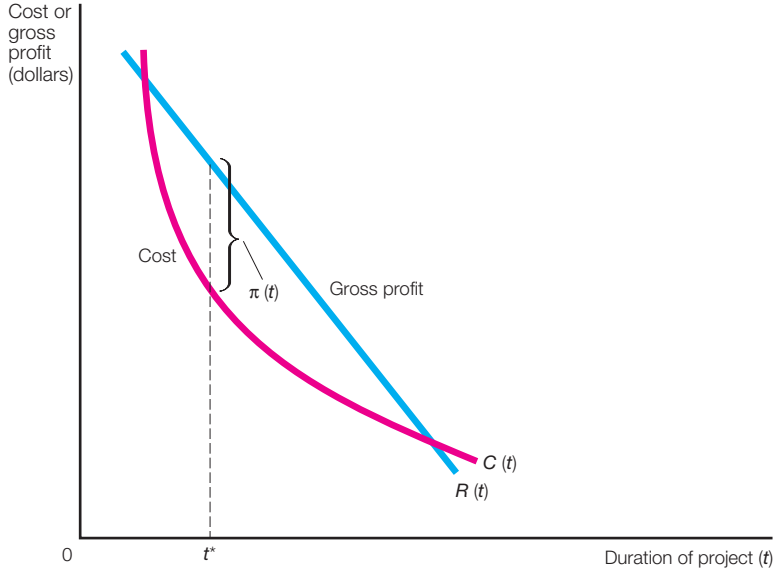
TIME-COST TRADE-OFFS

For a particular innovator, there is likely to be a time-cost trade-off function, like that in Figure A.4. If the firm cuts the total time taken to develop and introduce the innovation, it incurs higher costs. As the development schedule is shortened, more tasks must be carried out concurrently rather than sequentially, and since each task provides information useful in carrying out the others, there are more false starts and wasted designs. Also, diminishing returns set in as more and more technical workers are assigned simultaneously to the project.

FIGURE A.4

Time-Cost Trade-off Function and Optimal Duration of the Project

The optimal duration of the project is t^* years.



Faced with this time-cost trade-off function, how quickly should the firm develop and introduce the innovation? Clearly, the answer depends on the relationship between the present value of profit (gross of innovation cost) from the innovation and how quickly the firm develops and introduces it. (For a detailed discussion of the concept of present value, see Appendix C.) If $R(t)$ is the present value of gross profit if the duration of the project is t years and the time-cost trade-off function is $C(t)$, profit equals

$$\pi(t) = R(t) - C(t) \quad (\text{A.7})$$

and the first-order condition for profit maximization is

$$\frac{\Delta C}{\Delta t} = \frac{\Delta R}{\Delta t} \quad (\text{A.8})$$

In Figure A.4, the optimal duration of the project is t^* years, since $\pi(t)$, which is the vertical difference between $R(t)$ and $C(t)$, is greatest when this is the duration of the project.

To illustrate, consider the Hanover Company, which wants to develop a new plastic. Its vice president for research and development believes that the time-cost trade-off function for this project is

$$C = 520 - 100t + 5t^2$$

where C is cost (in thousands of dollars) and t is the duration of the project (in years). This equation assumes that $t \geq 1$, since it is believed that the project cannot be carried out in less than a year. Hanover's president believes that

$$R = 480 - 20t$$

where R is the present value of profit (gross of innovation cost) from the innovation (in thousands of dollars). Since

$$\frac{\Delta C}{\Delta t} = \frac{\Delta(520 - 100t + 5t^2)}{\Delta t} = -100 + 10t$$

$$\frac{\Delta R}{\Delta t} = \frac{\Delta(480 - 20t)}{\Delta t} = -20$$

it follows from equation (A.8) that the firm should choose t so that

$$\begin{aligned} -100 + 10t &= -20 \\ t &= 8 \end{aligned}$$

QUANT OPTION

The profit for period t is

$$\pi(t) = R(t) - C(t)$$

The first order condition for profit maximization is

$$\frac{d\pi(t)}{dt} = \frac{dR(t)}{dt} - \frac{dC(t)}{dt} = 0$$

or

$$\frac{dR}{dt} = \frac{dC}{dt}$$

If

$$C = 520 - 100t + 5t^2$$

and

$$R = 480 - 20t$$

then

$$\frac{dC}{dt} = -100 + 10t$$

and

$$\frac{dR}{dt} = -20$$

then

$$\frac{dC}{dt} = -100 + 10t = -20 = \frac{dR}{dt}$$

or

$$10t = 80$$

or

$$t = 8$$

In other words, the Hanover Company should carry out the project in about eight years.

THE LEARNING CURVE

In many industries, technological change is due in considerable part to the learning and on-the-job experience that occurs as a firm produces more and more of a given item. Therefore, holding the firm's output rate constant, its average cost declines with increases in its cumulative total output (that is, the total number of items of this sort that it has produced in the past). For example, production of the first 100 machine tools of a particular type may require about 50% more hours of labor than production of the second 100 machine tools of this type, even though the number of machine tools produced per month remains about the same. Thus, the average cost of this machine tool falls substantially as cumulative total output grows.

One should distinguish between cost reductions due to learning and those due to greater current output. Holding constant the number of these machine tools produced by this firm in the past, it is quite possible that the average cost of producing such a machine tool during the current period declines as more of them are produced. But, this is different from learning. Holding constant the number of such machine tools produced currently, if the average cost is inversely related to the firm's previous total output of this machine tool, this is due to learning.

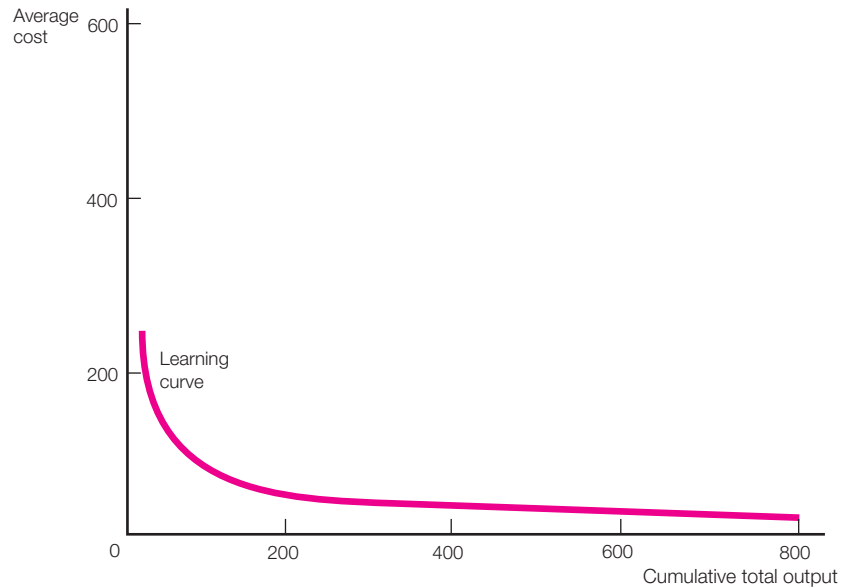
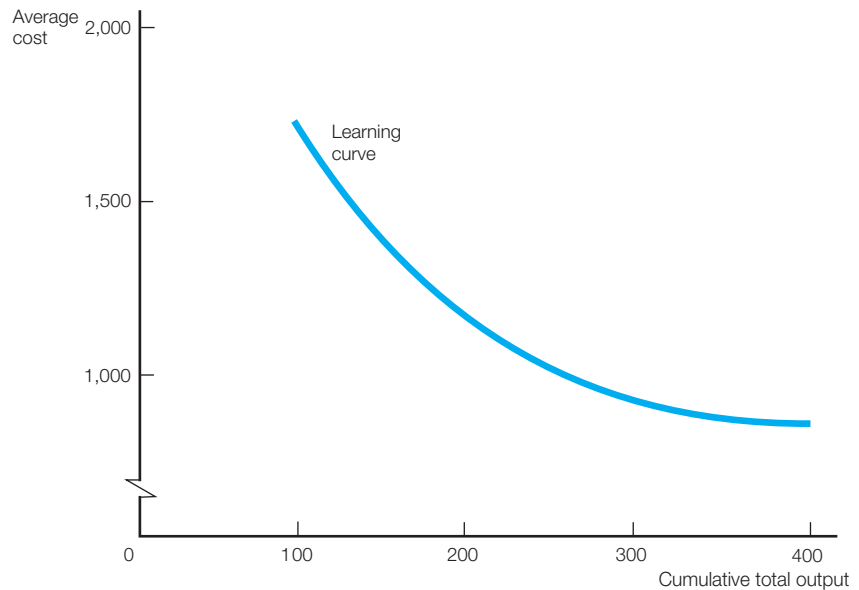
Managers, economists, and engineers often use the learning curve to represent the extent to which the average cost of producing an item falls in response to increases in its cumulative total output. Figure A.5 shows the learning curves for two actual products: a piece of optical equipment (produced by Optical Equipment Company) and a portable turbine (produced by Solar International, Inc.). As you can see, learning results in major reductions in the average cost of both products. Of course, these cost reductions are not automatic: They occur only if workers and managers strive for increased efficiency. But for many products of this sort, a doubling of cumulative output tends to produce about a 20 or 30% reduction in average cost.

APPLICATIONS OF THE LEARNING CURVE

Many firms adopted pricing strategies based on the learning curve. Consider the case of Texas Instruments, a major producer of semiconductor chips and other electronic products. When the semiconductor industry was relatively young, Texas Instruments priced its product at less than its then-current average costs to increase its output rate and its cumulative total output. Believing that the learning curve was relatively steep, it hoped that this would reduce its average costs to such an extent that its product would be profitable to produce and sell at this low price.

FIGURE A.5**Learning Curves**

Average cost declines with increases in cumulative total output.

**A.** Optical equipment produced by Optical Equipment Company**B.** Portable turbine produced by Solar International, Inc.

This strategy was extremely successful. As Texas Instruments continued to cut its price, its rivals began to withdraw from the market, its output continued to go up, its costs were reduced further, and its profits rose.⁶

The learning curve is expressed as

$$C = aQ^b \quad (\text{A.9})$$

where C is the input cost of the Q th unit of output produced. If this relationship holds exactly, a is the cost of the first unit produced. The value of b is negative, since increases in cumulative total output reduce cost. If the absolute value of b is large, cost falls more rapidly with increases in cumulative total output than it would if the absolute value of b were small. Taking logarithms of both sides of equation (A.9)

$$\log C = \log a + b \log Q \quad (\text{A.10})$$

In this logarithmic form, b is the slope of the learning curve.

To estimate the learning curve from historical data concerning cost and cumulative output, one can use the regression techniques in Chapter 4. As shown in equation (A.10), $\log C$ is a linear function of $\log Q$. Therefore, to estimate a and b , we can regress $\log C$ on $\log Q$. (In other words, $\log C$ is the equivalent of Y in Chapter 4, and $\log Q$ is the equivalent of X .) Of course, the values of a and b vary from product to product and firm to firm.

To illustrate how the learning curve can be used in specific cases, suppose that the controller of the Killian Company, a maker of a particular type of machine tool, finds that, for her firm, the learning curve (in logarithmic form) is

$$\log C = 4.0 - 0.30 \log Q$$

where C is expressed in dollars. (That is, $\log a = 4.0$ and $b = -0.30$.) From this equation, she can estimate how much the cost per unit will go down in the future. For example, if she wants to estimate the cost of the 100th machine tool of a particular type, the answer is

$$\log C = 4.0 - 0.30 \log 100 = 4.0 - 0.30(2) = 3.4$$

Since the antilog of 3.4 is 2,512, the answer is that the cost will be \$2,512.

HENRY FORD'S MODEL T AND DOUGLAS AIRCRAFT'S DC-9

The learning curve is nothing new. Between 1908 and 1923, the price of Henry Ford's famous Model T automobile fell from over \$3,000 to under \$1,000, owing in considerable measure to cost reductions due to learning. Ford worked hard to push costs down in this way. Standardization was increased. His product line was less diverse than those of his competitors, and model improvements occurred

6. For a classic paper concerning learning curves, see K. Arrow, "The Economic Implications of Learning by Doing," *Review of Economic Studies*, June 1962, Vol. 29(3), pp. 155–173. The Boston Consulting Group was a leading advocate of its application to corporate planning.

less frequently. The production throughput time was reduced, and the division of labor was increased.

However, not all firms have been as successful as the Ford Motor Company in reducing costs in this way. In cases in which labor turnover is high or a firm cannot obtain workers with the necessary skills, expected cost reductions due to learning may not materialize. For example, when Douglas Aircraft planned the production of the DC-9 airframe, it anticipated little problem in getting qualified workers. But when the time came, the labor market was so tight in Los Angeles that Douglas soon lost 12,000 of the 35,000 workers it hired. The result was that, contrary to the firm's expectations, costs did not fall as a result of learning, substantial losses were incurred, and the firm was forced into a merger (resulting in McDonnell Douglas, which is now part of Boeing).⁷

DIFFUSION MODELS

Another type of technological forecasting technique is based on the use of econometric diffusion models, which analyze the rate at which an innovation spreads. Although these models forecast the diffusion of new processes and products already in existence rather than the occurrence of future inventions, this limitation may be less important than it seems, since the inventions that already occurred are sometimes all that really matter in the short and the intermediate runs. In part, this is because it frequently takes a long time for an invention to be commercially introduced. For example, it took about nine years before catalytic cracking, a major innovation in oil refining, was first used.

The diffusion process, like the earlier stages in the creation and assimilation of new processes and products, is essentially a learning process. However, rather than being confined to a research laboratory or a few firms, the learning takes place among a considerable number of users and producers. When the innovation first appears, potential users are uncertain of its nature and effectiveness, and they tend to view its purchase as an experiment. Sometimes, considerable additional research and development is required before the innovation is successful; sometimes, despite attempts at redesign and improvement, the innovation never is a success. Information regarding the existence, characteristics, and availability of the innovation is circulated by the producers through advertisements and sales representatives; information regarding the reaction of users to the innovation tends to be circulated informally and through the trade press.

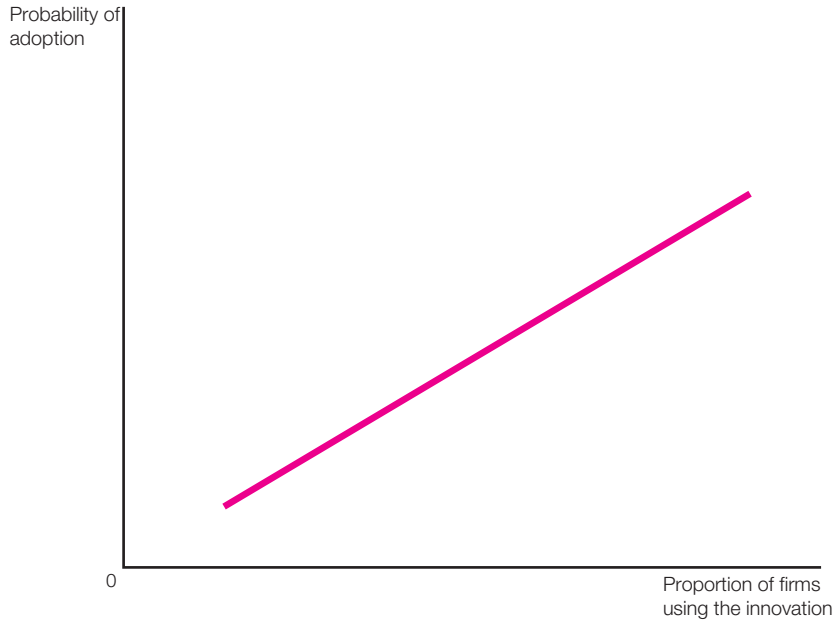
Figure A.6 illustrates an important aspect of the process by which new techniques spread throughout an industry. The figure shows the probability that a firm not using an innovation will adopt it in the next few months and is influenced by the proportion of firms in the industry already using it. Specifically, as the number of firms adopting an innovation increases, the probability of its adoption by a nonuser increases. This is because the risks associated with its introduction grow

7. J. Macklin, "Douglas Aircraft's Stormy Flight Path," *Fortune*, December 1966.

FIGURE A.6

Relation between Probability of a Nonuser's Adopting a Process Innovation and the Proportion of Firms Already Using the Innovation

This relationship tends to be direct.



smaller, competitive pressures mount, and bandwagon effects increase as experience and information regarding an innovation accumulate.

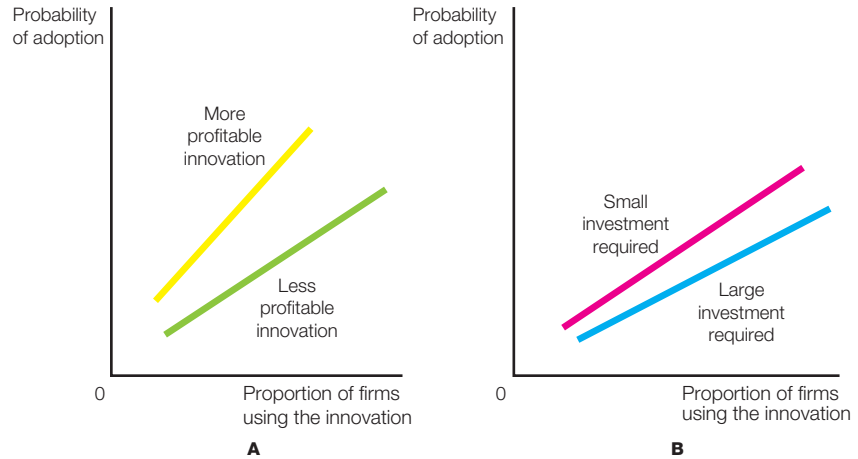
Other important aspects of the diffusion process are brought out by Figure A.7. Panel A shows that the probability that a nonuser will adopt the innovation is higher for more profitable innovations than for less profitable innovations, holding constant the proportion of firms in the industry already using it. The more profitable the investment in an innovation promises to be, the greater is the probability that a firm's estimate of its potential profitability compensates for the risks involved in its installation.

Panel B of Figure A.7 shows that the probability a nonuser will adopt the innovation is higher for innovations requiring fairly small investments, holding constant the proportion of firms in the industry that are already using it (and the profitability of the innovation). This is because firms are more cautious before committing themselves to large, expensive projects; and they have more difficulty in financing them.

FIGURE A.7

Effect of Profitability of the Innovation and the Size of Investment Required to Adopt the Innovation on Probability of Adoption

This probability tends to be directly related to profitability and inversely related to the size of the investment.



If the relationship in Figure A.6 holds, it can be shown that $P(t)$, the proportion of firms using the innovation, increases in accord with the S-shaped growth curve shown in Figure A.8. The formula for this growth curve (often called the *logistic curve*) is

$$P(t) = \frac{1}{1 + e^{-(A+Bt)}} \quad (\text{A.11})$$

where A and B are parameters that vary from innovation to innovation and t represents time. Whether the diffusion process goes on slowly, as in curve L in Figure A.8, or quickly, as in curve M , depends on the profitability of the innovation and the size of investment it requires. This model has much in common with the models used by epidemiologists to represent the spread of contagious diseases. Firms in a wide variety of industries have found that it can explain reasonably well the available data concerning the diffusion process.⁸

FORECASTING THE RATE OF DIFFUSION OF NUMERICALLY CONTROLLED MACHINE TOOLS

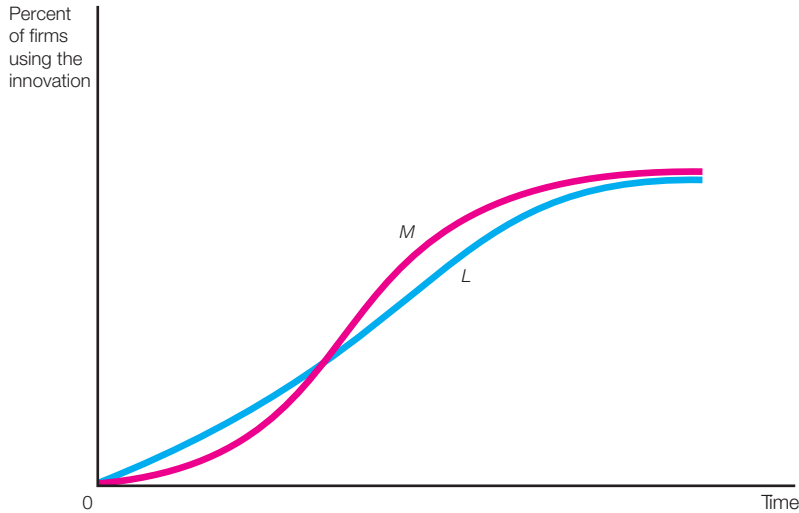
To illustrate the use of diffusion models for forecasting, consider a study carried out for the Small Business Administration to forecast the percentage of firms in the tool and die industry that would be using numerically controlled machine

8. See E. Mansfield, *Industrial Research and Technological Innovation* (New York: Norton, 1968); E. Mansfield et al., *The Production and Application of New Industrial Technology* (New York: Norton, 1977); V. Mahajan and Y. Wind, eds., *Innovation Diffusion Models of New Product Acceptance* (Cambridge, MA: Ballinger, 1986); and E. Mansfield, "Contributions of New Technology to the Economy," in *Technology, R and D, and the Economy*, ed. Bruce Smith and Claude Barfield (Washington, DC: Brookings Institution and American Enterprise Institute, 1996).

FIGURE A.8

Growth over a Period of Time in the Proportion of Firms Using the Innovation

Both growth curves L and M are S-shaped.



tools two years after the time of the study.⁹ When the study was carried out, about 20% of the firms in the National Tool, Die, and Precision Machining Association were using numerically controlled machine tools. To use the model described in the previous section, data were obtained, both from a mail survey and an interview study, of the past growth over time in the percentage of tool and die firms using such machine tools. Based on these data, and using the regression techniques in Chapter 4, estimates of A and B in equation (A.11) were made. To see how these estimates were calculated, note that equation (A.11) implies that

$$\ln \{P(t)/[1 - P(t)]\} = A + Bt \quad (\text{A.12})$$

Thus, A and B can be estimated by regressing $\ln \{P(t)/[1 - P(t)]\}$ on t .¹⁰ (The natural logarithm of any number, say, Y , is designated $\ln Y$.)

Given estimates of A and B , equation (A.11) can be used to forecast $P(t)$ for future values of t . On the basis of the interview data, the model forecasted that about 33% of the firms would be using numerically controlled machine tools. On the basis of mail-survey data, the model forecasted that about 37% of the firms would be using them.

To see how these forecasts compare with those obtained on the basis of other methods, two alternative types of forecasts were made. First, the firms—both in interviews with a carefully selected sample of industry executives and in the

9. See Mansfield et al., *The Production and Application of New Industrial Technology*.

10. This is only a rough estimation technique, but it is adequate for present purposes.

PROBLEM SOLVED: Information Technology and Productivity Growth

The U.S. Bureau of Labor Statistics reports the annual output per worker in the United States increased by 4.8% in 2002, and over 5% in 2003. During the 1980s and 1990s the annual growth in output per worker was less than 2.7%. While many articles attribute this productivity increase to the expanded use of information technology (IT), recent research shows that technology alone only accounts for a small percentage of the increase. More important than technology is how managers re-design their work processes around the technology.

Professors Eric Brynjolfsson of MIT and Lorin Hitt of the The Wharton School studied over 1,100 large U.S. corporations. They concluded the critical question facing managers should not be, "Does IT pay off," but "How can we best use computers?" They found that complementary organizational capital assets coupled with IT provide most of the increases in productivity. These complementary assets include human capital, training, work processes and routines, knowledge transfer, and corporate culture and value. For every dollar invested in IT hardware, companies spend up to nine dollars on these complementary assets. Their conclusion is that IT does not in itself produce productivity gains, rather it is the management around the use of IT that produces the gains.

For example, they examined the use of IT in both Wal-Mart and K-Mart. They found that while Wal-Mart did have a greater use of IT per employee, the bigger differences were in how the organization is designed around IT. Decision making in Wal-Mart is more decentralized, there is a greater use of teams, and compensation is more performance based.

Technological advances like electric motors and robotics essentially replaced the brawn of humans. Computers are different since they do not replace the decision-making abilities of humans, rather they complement those abilities. In fact, studies find companies that use IT extensively employ workers who are better trained, educated, and skilled. These companies leverage this combination of computers and cognitive power by empowering employees through the use of intranets and database sharing. They also motivate their employees to make good decisions by structuring performance based incentive systems.

While IT itself holds the potential for increased productivity within a firm, this potential is only realized when managers take the appropriate actions to re-design organizational assets to complement the technology. Brynjolfsson and Hitt suggest a set of organizational design changes to enhance the productivity of IT systems. These include:

- automation of routine tasks
- use of highly skilled employees
- decentralized decision making
- more efficient information flow
- greater use of performance based incentives

Sources: Steve Lohr. "Technology and Worker Efficiency," *New York Times*. February 2, 2004 at www.nytimes.com/2004/02/02/technology/02new.html; Erik Brynjolfsson, "The IT Productivity Gap," ebusiness.mit.edu/erik/Optimize/pr_roi.html.

mail survey of the industry—were asked whether they planned to begin using numerical control in the next two years. Since considerable lead time is required in obtaining numerical control, it seemed reasonable to suppose that their replies would have some forecasting value. The results of the interviews indicated that

TABLE A.2

Alternative Two-Year Forecasts of the Percentage of Firms in the U.S. Tool and Die Industry Using Numerical Control, and the Actual Percentage

Type of forecast	Based on interview data	Based on mail survey
Model	33	37
Plans of tool and die firms	33	43
Median forecast by machine tool builders	30	30
Actual percentage	37	37

about 16% of nonusers planned to use numerical control; the results of the mail survey indicated that this was the case for 28% of the nonusers. Therefore, the forecast was 33%, based on the interview data, or 43%, based on the mail survey.

Second, forecasts were obtained from the machine tool builders, the firms presumably closest to and best informed about the market for numerically controlled machine tools. About 25 of the 150 members of the National Machine Tool Builders Association provided forecasts. The results showed a considerable amount of variation, but the median forecast was about 30%.

How accurate were these forecasts? Which forecasting approach was most accurate? Table A.2 shows that the model's forecast based on the data from the mail survey was almost precisely correct and that the model's forecast based on the interview data was off by only 4 percentage points. Regardless of whether we look at results based on the interview data or the mail survey data, the model forecasts better than the other two techniques. Moreover, it forecasts better than simple extrapolation by "naive" models.¹¹ Certainly, this is encouraging. On the basis of these and other results, it appears that this simple model may be of use in forecasting the rate of diffusion. Of course, this does not mean that it is anything more than a crude device or that it can be applied in situations in which its basic assumptions do not hold. But it does mean that, used with caution, the model may perform at least as well as other commonly used forecasting devices.

SUMMARY

1. Technological change is the advance of technology; it often results in a change in the production function for an existing product or in a new product. The rate of technological change is often measured by changes in productivity. Changes in total factor productivity are often used by firms to measure changes in efficiency.

11. Specifically, the model forecast better than naive models that assumed that the increase in the percentage of firms using numerical control would be the same amount, in absolute or relative terms, during the next two years as it had been during the previous two years.

2. Research and development can be regarded as a process of uncertainty reduction, or learning. Chance plays a large role in research and development, and many projects use parallel efforts to help cope with uncertainty. Techniques are presented in this chapter to indicate when parallel efforts should be used.

3. An R and D project's likelihood of economic success is the product of three factors: the probability of technical success, the probability of commercialization (given technical success), and the probability of economic success (given commercialization). All three seem to be directly related to how quickly an R and D project is evaluated for its economic, as opposed to only technical, potential.

4. To promote successful R and D, there must be a strong link between R and D and marketing personnel and project selection techniques must be effective. However, this does not mean that more complicated quantitative selection techniques need be used.

5. For a particular innovation, there is likely to be a time-cost trade-off function. If the firm cuts the total time taken to develop and introduce the innovation, it incurs higher costs. Time-cost trade-off functions vary from firm to firm, because some firms are more adept and experienced than others in developing and introducing a particular innovation. The optimal duration of the project is the time interval where the discounted gross profits exceed the discounted cost by the maximum amount.

6. In many industries, there is a learning curve, which shows the extent to which the average cost of producing an item falls in response to increases in its cumulative total output. This learning curve plays an important role in pricing. For example, Texas Instruments successfully priced its product at less than its then-current average cost to move quickly down the learning curve. Regression techniques can be applied to estimate the learning curve for a particular product.

7. As the number of firms adopting a new process increases, the probability of its adoption by a nonuser increases. Also, the probability that a nonuser will adopt the innovation is higher for more-profitable innovations than for less-profitable innovations and for innovations requiring small investments than for those requiring large investments. A model based on these propositions can sometimes be of use in forecasting the rate of diffusion of an innovation.

PROBLEMS

1. The Monroe Corporation uses three inputs: labor, energy, and materials. In 2011, it used 20,000 hours of labor, 50,000 kilowatt-hours of energy, and 10,000 pounds of materials to produce 200,000 pounds of output. In 2012, it used 30,000 hours of labor, 100,000 kilowatt-hours of energy, and 14,000 pounds of materials to produce 300,000 pounds of output. In 2011, the price of labor was \$10 per hour, the price of a kilowatt-hour of energy was \$0.02, and the price of a pound of materials was \$5.

- a. What was the total factor productivity in 2011?
 - b. What was the total factor productivity in 2012?
 - c. What is the base year in the preceding calculations?
2. The chief scientist at the Roosevelt Laboratories estimates that the cost (in millions of dollars) of developing and introducing a new type of antiulcer drug equals

$$C = 100 - 19t + 0.5t^2, \quad \text{for } 1 \leq t \leq 6$$

where t is the number of years taken to develop and introduce the new drug. The discounted profit (gross of innovation cost) from a new drug of this type (in millions of dollars) is estimated to equal

$$R = 110 - 15t, \quad \text{for } 1 \leq t \leq 6$$

- a. The managers of the Roosevelt Laboratories are committed to developing and introducing this new drug within six years, and it is impossible to develop and introduce it in less than one year. What project duration would minimize cost?
 - b. Why does R decline as t increases?
 - c. What is the optimal project duration? Why?
3. The Flynn Company produces a particular type of commercial truck. Its chief engineer regresses the logarithm of the input cost of the Q th truck produced on the logarithm of Q , the result being

$$\log C = 5.1 - 0.25 \log Q$$

where C is input cost (in dollars).

- a. What is the estimated input cost of the 100th truck produced?
 - b. What is the estimated input cost of the 200th truck produced?
 - c. By what percentage does unit input cost decline if output is doubled (from 100 to 200 trucks)?
4. The Martin Company's president wants to estimate the proportion of chemical firms that will be using a particular new process in 2012. One of her assistants regresses $\ln \{m(t)/[n - m(t)]\}$ on t where $m(t)$ is the number of chemical firms using this process in year t and n is the total number of chemical firms that can use this process. Measuring t in years from 1994, the regression is

$$\ln \left[\frac{m(t)}{n - m(t)} \right] = -4.0 + 0.22t$$

- a. Prove that, if the proportion of chemical firms using the new process increases in accord with the logistic curve in equation (A.11), $\ln \{m(t)/[n - m(t)]\}$ is a linear function of t .
- b. On the basis of the preceding regression, can you estimate A and B (the parameters of the logistic curve in equation (A.11))? If so, how?

- c. Forecast the percentage of chemical firms using the new process in 2012.
5. In the aircraft industry, many studies indicate that a doubling of cumulative output results in about a 20% reduction in cost. If the cost of the 30th unit produced of a particular aircraft is \$12 million, what is the cost of the 60th unit produced? Of the 120th unit produced?
6. The Bureau of Labor Statistics produced data showing that output per hour of labor in blast furnaces using the most up-to-date techniques has sometimes been about twice as large as the industry average.
 - a. How can such large differences exist at a given time? Why don't all firms continually adopt the most up-to-date techniques?
 - b. Should firms always adopt techniques that maximize output per hour of labor? Why or why not?
 - c. Should firms adopt techniques that maximize output per dollar of capital? Why or why not?
7. The Russell Corporation is trying to develop an engine that will emit fewer pollutants. There are two possible approaches to this technical problem. If either one is adopted, there is a 50–50 chance that it will cost \$2 million to develop the engine and a 50–50 chance that it will cost \$1 million to do so.
 - a. If the firm chooses one of the approaches and carries it to completion, what is the expected cost of developing the engine?
 - b. If the two approaches are run in parallel and the true cost of development using each approach can be determined after \$150,000 has been spent on each approach, what is the expected cost of developing the engine? (Note that the total cost figure for each approach, if adopted, includes the \$150,000.)
 - c. Should parallel approaches be used?
8. To help decide whether particular R and D projects should be carried out, some firms compare the estimated cost of each project with the estimated profits it will earn. To carry out such an analysis, the firm's personnel must estimate how much the R and D project would cost if it were carried out. In one major drug firm, the frequency distribution of 49 projects by the ratio of actual to estimated cost is as follows:

Actual cost divided by estimated cost	Number of projects
Less than 1.01	6
1.01 and under 2.01	24
2.01 and under 3.01	16
3.01 and under 4.01	3

- a. If this firm were using this technique to help determine whether particular R and D projects should be carried out, what problems would be encountered?
 - b. How might the firm try to cope with these problems?
9. The Monroe Corporation wants to develop a new process that would reduce its costs by 10%. There are two ways to go about developing such a process. If the first way is adopted, there is a 0.6 probability that it will cost \$5 million to develop the process and a 0.4 probability that it will cost \$3 million to do so. If the second way is adopted, there is a 0.7 probability that it will cost \$3 million and a 0.3 probability that it will cost \$5 million.
- a. If the first way is adopted, what is the expected cost of developing the new process?
 - b. If the second way is adopted, what is the expected cost?
 - c. If the two approaches can be run in parallel and the true cost of development using each approach can be determined after \$500,000 has been spent on each approach, what is the expected cost? (Assume that the outcomes of the two approaches are independent. Also, note that the total cost figure for each approach, if adopted, includes the \$500,000.)
10. On the basis of past growth of the percentage of firms in the machinery industry using robots, this percentage can be approximated by

$$P(t) = \frac{1}{1 + e^{-(6.1 + 0.41t)}}$$

where $P(t)$ is this percentage and t is measured in years from 1990.

- a. During which year did about 25% of the firms in the machinery industry use robots?
- b. During which year did 50% of the firms in this industry use robots?

BUSINESS AND ECONOMIC FORECASTING

Many corporations have corporate economists or hire consulting firms to forecast sales or profits. These forecasts can be done on a short-term basis, for example, next quarter, or on a longer-term basis, for the next year, five years, or ten years. Security analysts forecast quarterly and yearly earnings, and one can hear their consensus via First Call before earnings are about to be released. Popular business press publications, like *Bloomberg BusinessWeek* and the *Wall Street Journal*, offer similar results. The Federal Reserve and the Department of Commerce also conduct business forecasts as do trade associations for their industries.

Managers of all sorts are routinely involved in decision making that is informed by forecasts. Sales forecasts, for instance, may dictate raw material orders, production run schedules, and hiring decisions. An organization's mid-level managers are likely to make these decisions. On the other hand, long-term growth forecasts for an industry may entail capital expansion or disinvestment of assets, decisions likely to be made by higher-level managers within the organization, most likely with the board of directors' approval.

In this appendix, we take up the techniques used by many business and economic forecasters. As you work through these subjects, keep in mind that economic forecasting is not an exact science. Too many variables exist to precisely model the economic system or even a small part of it. Even so, most corporations

LEARNING OBJECTIVES

- Survey Techniques
- Taking Apart a Time Series
- How to Estimate a Linear Trend
- How to Estimate a Nonlinear Trend
- Seasonal Variation
- Calculation of Seasonal Variation
- Cyclical Variation
- Elementary Forecasting Techniques
- How Leading Indicators Are Used
- How Econometric Models Are Used
- The Purvere Corporation: A Numerical Example
- "Study Your Residuals"
- Summary
- Problems
- Appendix: Exponential Smoothing and Forecasting

prefer rigorous data to intuition or hunch (and, in fact, managers have a fiduciary responsibility to the owners of the firm to practice due diligence in following rational procedures for decision making).

As we shall see, regression plays a major role in many of these forecasting techniques, including the econometric models that are the staple of the leading private and public economic forecasters.

SURVEY TECHNIQUES

One of the simplest forecasting devices is to survey firms or individuals to determine what they believe will occur. Consider the surveys carried out to forecast firms' expenditures. For example, the U.S. Department of Commerce and the Securities and Exchange Commission conduct surveys of business intentions to buy plant and equipment. Still other surveys are aimed at measuring consumer intentions. The Survey Research Center at the University of Michigan and other such groups provide information on planned purchases of automobiles, appliances, and housing. Also, they indicate the extent of consumer confidence in the economy, which is an important factor influencing consumers' spending decisions. Surveys of this type are of value in forecasting the sales of many products. They provide a wealth of information to the forecaster.

At least two types of information can be obtained from surveys. First, they can provide us with the respondent's forecast of some variable over which he or she has no control. For example, the University of Michigan obtains data from consumers concerning their forecasts of the rate of inflation. Second, surveys can provide us information concerning what people or firms believe they will do. For example, the National Federation of Independent Business surveys firms to determine whether, and to what extent, they plan to increase their prices.

Suppose a survey is used to forecast some variable, such as the sales of a particular firm. How can we determine how reliable this forecasting technique seems to be? One commonly used measure of the size of the forecast error is the root-mean-squared forecast error, which is defined as

$$E = \left(\sum_{i=1}^n (Y_i - F_i)^2 / n \right)^{0.5}$$

where F_i is the i th forecast, Y_i is the corresponding actual value, and n is the number of forecasts for which we have data concerning the size of the forecast errors. Therefore, if the forecasts for 2011, 2012, and 2013 are \$110 million, \$120 million, and \$130 million, and if the actual values are \$105 million, \$122 million, and \$127 million, respectively, the root-mean-squared forecast error equals

$$\sqrt{\frac{(105 - 110)^2 + (122 - 120)^2 + (127 - 130)^2}{3}} = 3.56$$

or \$3.56 million dollars. This measure of forecast error is used to evaluate forecasts, no matter whether they are based on surveys or other techniques. Clearly, the lower is the root-mean-squared forecast error, the better the forecasting technique.

TAKING APART A TIME SERIES

Although surveys are of considerable use, most major firms seem to base their forecasts in large part on the quantitative analysis of economic time series. The classic approach to economic forecasting, devised primarily by economic statisticians, was essentially descriptive. It assumed that an economic time series could be decomposed into four components: trend, seasonal variation, cyclical variation, and irregular movements. More specifically, it assumed that the value of an economic variable at a certain time could be represented as the product of each of these four components. For example, the value of a company's sales in January 2011 was viewed as equal to

$$Y = T \times S \times C \times I \quad (\text{B.1})$$

where T is the trend value of the firm's sales during that month, S is the seasonal variation attributable to January, C is the cyclical variation occurring that month, and I is the irregular variation that occurred then.¹ Each of these components is defined next.

Trend

A trend is a relatively smooth long-term movement of a time series. For instance, the civilian labor force of the United States increased rather steadily between 1948 and 2011, as shown in Figure B.1. Hence, there has been an upward trend in the U.S. civilian labor force. Of course, not all trends are upward. The trend in farm employment in the United States has generally been downward, as shown in Figure B.2.² Whether upward or downward, the trend of a time series is represented by a smooth curve. In equation (B.1), T is the value of the firm's sales predicted for January 2011, on the basis of such a curve.

Seasonal Variation

In a particular month, the value of an economic variable is likely to differ from what would be expected on the basis of its trend because of seasonal factors. For example, consider the sales of a firm that produces Christmas trees. Since the demand for Christmas trees is much higher in the winter than in the summer, one would expect that the monthly time series of the firm's sales would show a pronounced and predictable seasonal pattern. Specifically, sales each year would tend to be higher in December than during the rest of the year. As we shall see, it is possible to calculate *seasonal indexes* that estimate how much each month departs from what would be expected on the basis of its trend. In equation (B.1), we must multiply the trend value T by the seasonal index S to allow for the effect of this seasonal variation.

1. In some versions of this model, the components are added rather than multiplied. That is, it is assumed that

$$Y = T + S + C + I$$

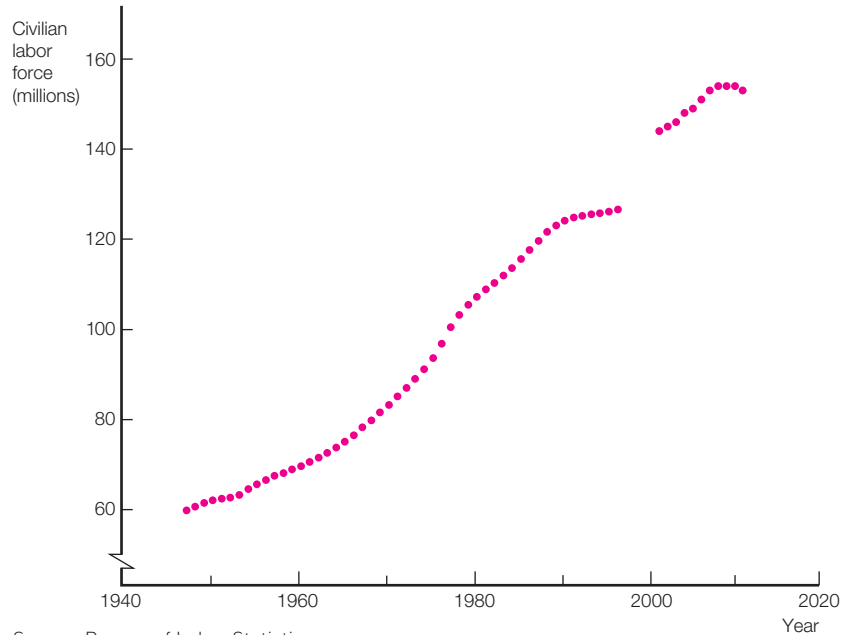
where Y is the value of the time series.

2. In still other cases, the trend is horizontal; that is, there is no upward or downward tendency in the time series. In these cases, it is often said that there is no trend.

FIGURE B.1

Civilian Labor Force of the United States, 1948–2011

This series exhibits a strong upward trend.



Source: Bureau of Labor Statistics.

Cyclical Variation

Another reason why an economic variable may differ from its trend value is that it may be influenced by the so-called business cycle. The general tempo of economic activity in our society has exhibited a cyclical nature, with booms being followed by recessions and recessions being followed by expansions. These cycles have not been regular or consistent (which is one reason why many economists prefer the term *business fluctuations* to *business cycles*); but unquestionably there has been a certain cyclical ebb and flow of economic activity, which has been reflected in a great many time series. For this reason, $T \times S$ is multiplied by C , which is supposed to indicate the effect of cyclical variation on the firm's sales in equation (B.1).

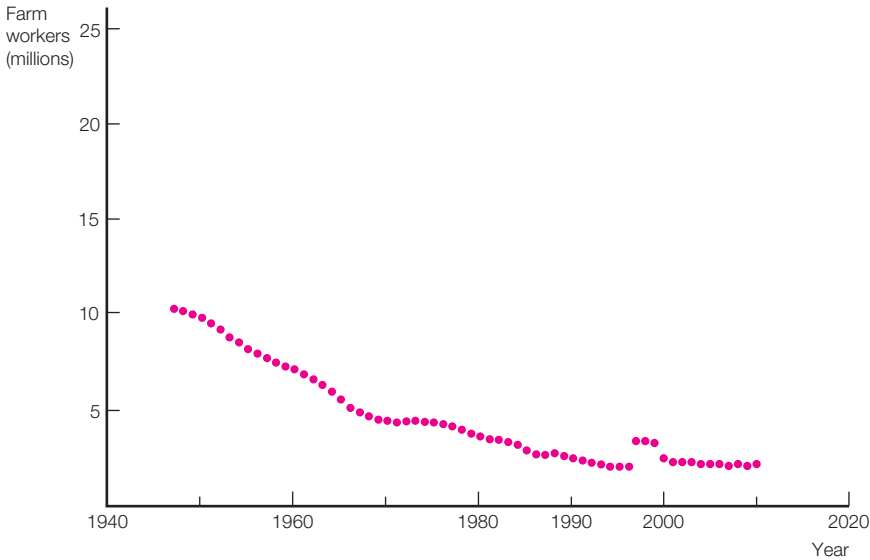
Irregular Variation

Once it has been multiplied by both S and C , the trend value T has been altered to reflect seasonal and cyclical forces. But in addition to these forces, *a variety of short-term, erratic forces is also at work.* Their effects are represented by I . Essen-

FIGURE B.2

Farm Employment in the United States, 1947–2011

This series exhibits a strong downward trend.



tially, I reflects the effects of all factors other than the trend, seasonal variation, and cyclical variation. According to the classic model, these irregular forces are too unpredictable to be useful for forecasting purposes.

HOW TO ESTIMATE A LINEAR TREND

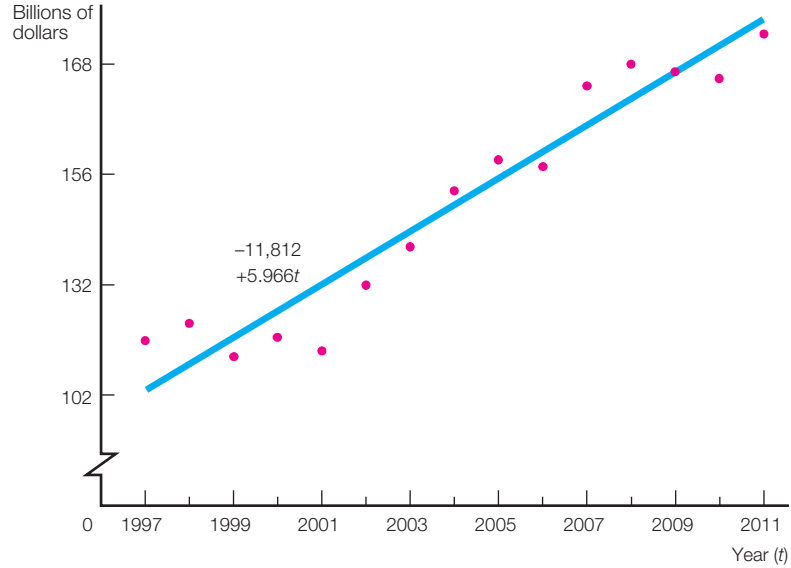
Managerial economists have carried out many studies to estimate the trend, seasonal variation, and cyclical variation in particular economic time series. In this and the following sections of this appendix, we encounter the methods used to estimate a trend; in subsequent sections, we take up seasonal and cyclical variation. First, we consider the case in which the long-term overall movement of the time series seems to be fairly close to linear. For example, this seems true for the sales of the ABC Corporation during the period 1997 to 2011. (These sales are plotted in Figure B.3.) In a case in which the trend seems to be linear, analysts frequently use the method of least squares to calculate the trend. In other words, they assume that, if the long-term forces underlying the trend were the only ones at work, the time series would be approximately linear. Specifically, they assume that

$$Y_t = A + Bt \quad (\text{B.2})$$

FIGURE B.3

Linear Trend in Sales, ABC Corporation, 1997–2011

The ABC Corporation's sales have risen rather steadily throughout the period.



The **trend value** is the value of Y_t that would result if only the trend existed

The **deviation from trend** is the difference of the actual value of Y_t and the value of Y_t assumed from equation (B.2).

where Y_t is the trend value of the variable at time t . (Note that t assumes values like 2009 or 2010 if time is measured in years.) The **trend value** is the value of the variable that would result if only the trend were at work. The deviation of Y , the actual value of the variable, from the trend value is the **deviation from trend**.

To illustrate the calculation of a linear trend, we examine the ABC Corporation's annual sales from 1997–2011. Since sales in year t is the dependent variable and t is the independent variable, it follows from our discussion in Chapter 4 that

$$b = \frac{\sum_{t=t_0}^{t_0+n-1} (S_t - \bar{S})(t - \bar{t})}{\sum_{t=t_0}^{t_0+n-1} (t - \bar{t})^2} \quad (\text{B.3})$$

and

$$a = \bar{S} - b\bar{t} \quad (\text{B.4})$$

where S_t is sales (in billions of dollars) in year t , t_0 is the earliest year in the time series (that is, 1997), $t_0 + n - 1$ is the latest year in the time series (that is, 2011), b is an estimate of B , and a is an estimate of A .

Inserting the data underlying Figure B.3 into equations (B.3) and (B.4), we find that the trend line is

$$S_t = -11,812 + 5.966t \quad (\text{B.5})$$

This trend line is plotted in Figure B.3.

HOW TO ESTIMATE A NONLINEAR TREND

Many time series do not exhibit linear trends. In some such cases, a quadratic function of time provides an adequate trend. Such a trend can be represented as

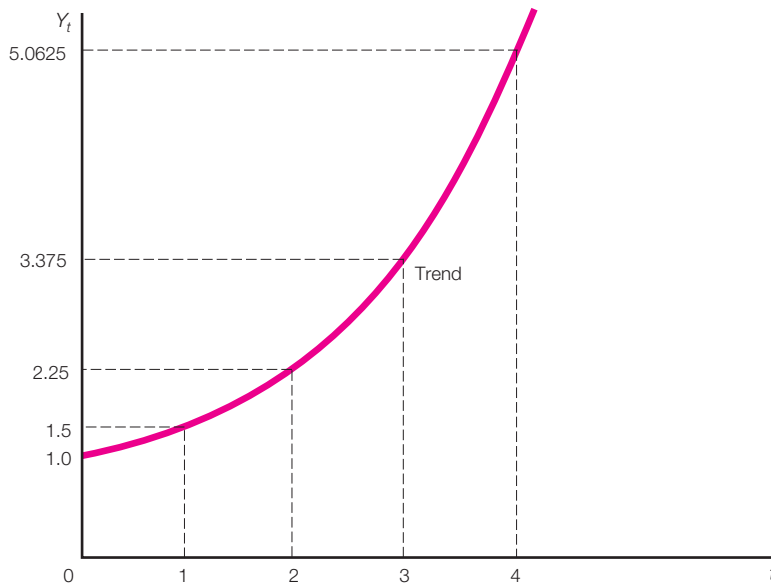
$$Y_t = A + B_1t + B_2t^2$$

To estimate A , B_1 , and B_2 , we can use the multiple regression techniques described in Chapter 4. As indicated there, standard computer programs are available to make these computations. The regression contains two independent variables: t and t^2 . Whether a quadratic trend is more appropriate than a linear trend can be determined by seeing whether it fits the data significantly better than a linear trend.

FIGURE B.4

Exponential Trend, Assuming $\beta = 1.5$, $\alpha = 1$

Many time series have exponential trends.



For many variables, an exponential curve provides a better-fitting trend than a quadratic curve. The equation for such a trend (shown in Figure B.4) is

$$Y_t = \alpha\beta^t \quad (\text{B.6})$$

where Y_t is the trend value of the time series at time t . A trend of this sort seems to fit many business and economic time series. It represents a situation in which the variable grows at a constant percentage rate per year. Therefore, if a firm's sales grow at about 5% per year, they are likely to exhibit an exponential trend.

If there is an exponential trend, we can take logarithms of both sides of equation (B.6)

$$\log Y_t = A + Bt \quad (\text{B.7})$$

where $A = \log \alpha$, and $B = \log \beta$. Since equation (B.7) is linear, we can estimate A and B by the method of least squares. Then, we can take antilogs of A and B to estimate α and β , the unknown coefficients in equation (B.6). (The average rate of increase of Y_t equals $\beta - 1$.)³ In this way, we can estimate the nonlinear trend shown in equation (B.6).

SEASONAL VARIATION

Many time series consist of monthly or quarterly rather than annual data. For such time series, managerial economists and decision makers must recognize that seasonal variation is likely to be present in the series. Seasonal variation in many economic time series is due to the weather. For example, sales of soft drinks are higher in the summer than in the winter. In other cases, such as sales of Christmas trees, seasonal variation is due to the location of a specific holiday (Christmas) on the calendar. Still other reasons for seasonal variation are the fact that some industries tend to grant vacations at a particular time of year, taxes have to be paid at certain times of the year, or schools tend to open at particular times of the year.

Managerial economists have devised methods for estimating the pattern of seasonal variation in a particular time series. In other words, they can determine the extent to which a particular month or quarter is likely to differ from what would be expected on the basis of the trend and cyclical variation in the same series. (In terms of the traditional model in equation (B.1), they can determine the value of S for each month or quarter.) For example, the marketing vice president for a manufacturer of soft drinks may tell the company's board of directors that U.S. production of soft drinks tends in June to be 5.9% higher than what the trend and cyclical variation in soft drink production would indicate. Or she may tell them that U.S. production of soft drinks in December tends to be 7.0% lower than the trend and cyclical variation would indicate.

3. If Y grows at a constant rate of 100r% per year,

$$Y_t = Y_0(1 + r)^t$$

where Y_0 is the value of Y in some base year (say 2010) and Y_t is its value t years after the base year. Taking logarithms of both sides of this equation,

$$\log Y_t = \log Y_0 + [\log(1 + r)]t$$

Therefore, $\log(1 + r)$ equals B , and the antilog of B (which is β) equals $(1 + r)$. Consequently, $r = \beta - 1$. In other words, as stated in the text, the average rate of increase of Y_t (which is r) equals $\beta - 1$. (Mathematical review: If X is the logarithm of Y , Y is called the antilog of X .)

TABLE B.1

Seasonal Variation in Production of Soft Drinks in the United States

Month	Seasonal index	Month	Seasonal index
January	93.4	July	112.4
February	89.3	August	113.4
March	90.7	September	108.3
April	94.9	October	103.9
May	99.0	November	95.8
June	105.9	December	93.0

The seasonal variation in a particular time series is described by a figure for each month, the **seasonal index**, which shows the way in which that month tends to depart from what would be expected on the basis of the trend and cyclical variation in the time series. For example, Table B.1 shows the seasonal variation in U.S. production of soft drinks. January's production tends to be about 93.4% of the amount expected on the basis of trend and cyclical variation, February's production tends to be about 89.3% of this amount, March's production tends to be about 90.7% of this amount, and so on. Figures of this sort can be used in a number of ways. One important application is to forecast what the time series will be in the future. For example, suppose that on the basis of the trend and cyclical variation, it appears likely that about 30 million gallons of soft drinks will be produced next January. If this is the case, a reasonable forecast of actual January production is $0.934(30 \text{ million}) = 28.02 \text{ million gallons}$, since January's production tends to be 93.4% of the amount expected based on trend and cyclical variation.

The **seasonal index** shows the deviation of a particular month's value from the value attributed to the trend and cyclical variation alone.

CALCULATION OF SEASONAL VARIATION

One way of calculating the seasonal variation in a time series is to use regression techniques. Suppose, for example, that a business analyst has a time series composed of quarterly values; that is, each observation pertains to the first, second, third, or fourth quarter of a year. If the analyst believes that the time series has a linear trend, he or she may assume that the value of the observation at time t equals

$$Y = A + B_1t + B_2Q_1 + B_3Q_2 + B_4Q_3 + e_t \quad (\text{B.8})$$

where Q_1 equals 1 if time t is the first quarter and 0 otherwise, Q_2 equals 1 if time t is the second quarter and 0 otherwise, Q_3 equals 1 if time t is the third quarter and 0 otherwise, and e_t is an error term.

The **expected value** of an observation is its long-term mean value.

It is important to understand the meaning of B_1 , B_2 , B_3 , and B_4 in equation (B.8). Clearly, B_1 is the slope of the linear trend, but what are B_2 , B_3 , and B_4 ? The answer is that B_2 is the difference between the expected value of an observation in the first quarter and the expected value of an observation in the fourth quarter when the effects of the trend are removed. (The **expected value** of an observation is its long-term mean value. To find its expected value, one multiplies each possible value of the observation by the probability of this value, and sums up the results.) To see that this is true, note that, if an observation pertains to time t , the first quarter of a particular year, its expected value equals

$$A + B_1t + B_2$$

according to equation (B.8). Similarly, if an observation pertains to time $t + 3$, the fourth quarter of the same year, its expected value equals

$$A + B_1(t + 3)$$

according to equation (B.8). Therefore, the difference between the expected value of an observation in the first quarter and the expected value of an observation in the fourth quarter equals

$$(A + B_1t + B_2) - [A + B_1(t + 3)] = B_2 - 3B_1$$

And, if we remove the effects of the trend (which is responsible for the last term on the right, $3B_1$), this difference equals B_2 ; this is what we set out to prove. When the effects of the trend are removed, one can show in the same way that B_3 is the difference between the expected value of an observation in the second quarter and the expected value of an observation in the fourth quarter, and B_4 is the difference between the expected value of an observation in the third quarter and the expected value of an observation in the fourth quarter.

Consequently, if equation (B.8) is valid, the analyst can represent the seasonal variation in the time series by the three numbers B_2 , B_3 , and B_4 . To estimate each of these numbers, ordinary multiple regression techniques can be used. The dependent variable is Y , and the independent variables are t , Q_1 , Q_2 , and Q_3 . The last three independent variables (Q_1 , Q_2 , and Q_3) are dummy variables. (A dummy variable is a variable that can assume only two values: 0 or 1.) Using the regression methods described in Chapter 4, the constants in equation (B.8) (A , B_1 , B_2 , B_3 , and B_4) can be estimated by the ordinary least-squares technique.

When using this procedure, the analyst assumes that seasonal effects are added to the trend value, as shown in equation (B.8). This differs from the traditional model in equation (B.1), where it is assumed that seasonal effects multiply the trend value (see footnote 1). The former assumption is appropriate in some cases, whereas the latter assumption is appropriate in others. Techniques based on both assumptions are useful.⁴

4. To calculate the seasonal variation based on the latter assumption, a four-quarter—or 12-month if the data are monthly—moving average can be used. For the details, see any business statistics book.

PROBLEM SOLVED: Forecasting the Demand for Blood Tests

North Carolina Memorial Hospital (now part of University of North Carolina Hospitals) has been interested in forecasting the number of blood tests it will perform. A simple model has been constructed that assumes that the number of tests per month increases according to a linear trend, and that the seasonal variation can be represented in the way described in equation (B.9). In other words, it is assumed that

$$Q = A + B_1t + B_2M_1 + B_3M_2 + \dots + B_{12}M_{11} + e_t,$$

where Q is the number of blood tests performed at the hospital in month t , M_1 equals 1 if month t is January and 0 otherwise, \dots , M_{11} equals 1 if month t is November and 0 otherwise, and e_t equals an error term. Therefore, B_2 is the difference between January and December in the expected number of tests, B_3 is the difference between February and December in the expected number of tests, and so on (when the effects of the trend are removed).

(a) Indicate how one can estimate the values of A , B_1 , B_2 , \dots , B_{12} . (b) Potential patients are reluctant to seek medical care during the Christmas holidays. Would you expect B_2 to be positive or negative? Why? (c) According to the hospital, the model forecasts "are being used to plan vacation schedules for employees

and to order supplies for the tests." Why would forecasts of this sort be useful for these purposes? (d) Forecasts based on this simple model have been reported to be "excellent." Forecasting errors have averaged only about 4.4%. On the other hand, forecasts based on exponential smoothing (a technique described in the appendix on page 819) did not perform so well. Do you think that a model of this sort will always outperform exponential smoothing?

SOLUTION (a) The values of these parameters can be estimated by calculating a multiple regression, where Q is the dependent variable and t , M_1 , M_2 , \dots , M_{11} are the independent variables. (b) Positive, because B_2 is the difference between January and December in the expected number of tests when the effects of the trend are removed. Because patients tend not to want such tests during the holidays, December would be expected to be below January in this regard. (c) If one can forecast the demand for blood tests, it is possible to estimate the number of employees and the quantity of supplies needed at various times. Clearly, this information is useful in scheduling vacations and purchases, among other things. (d) No. In some cases, one forecasting technique works well; in others, it works less well. No technique is universally better than the others discussed in this appendix.

To illustrate how this regression procedure can be used to estimate the seasonal variation in monthly data, suppose you have monthly data concerning the sales of a particular firm. If there is a linear trend, you can assume that

$$Y = A + B_1t + B_2M_1 + B_3M_2 + \dots + B_{12}M_{11} + e_t \quad (\text{B.9})$$

where Y is the firm's sales in month t , M_1 equals 1 if month t is January and 0 otherwise, M_2 equals 1 if month t is February and 0 otherwise, \dots , M_{11} equals 1 if month t is November and 0 otherwise, and e_t equals an error term. Using ordinary multiple regression techniques, you can estimate A , B_1 , B_2 , \dots , B_{12} . The

estimates of B_2, B_3, \dots, B_{11} , and B_{12} indicate the seasonal variation in the firm's sales. In particular, B_2 is the difference between January and December in the expected value of sales, B_3 is the difference between February and December in the expected value of sales, and so on, until B_{12} is the difference between November and December in the expected value of sales (when the effects of the trend are removed).

CYCLICAL VARIATION

Time series in business and economics frequently exhibit cyclical variation, such variation often being termed the **business cycle**. To illustrate what we mean by the business cycle or **business fluctuations**, we look at how national output has grown in the United States since 1980. Figure B.5 shows the behavior of gross domestic product (GDP) in constant dollars in the United States since 1980. Clearly, output has grown considerably during this period; indeed, GDP is more than three and a half times what it was 30 years ago. But this growth has not been steady. While the long-term trend has been upward, there have been periods, such as 1981–1982, 1990–1991, and 2008–2009 when national output declined.

The **full-employment level** of GDP is the total amount of goods and services that could have been produced if there had been full employment. Figure B.5 shows that national output tends to rise and approach (and perhaps exceed)⁵ its full-employment level for a while, then falters and falls below this level, then rises to approach it once more, then falls below it again, and so on. For example, output fell far below the full-employment level in the recession of 2008–2009 but was at this level when the economy was booming in the mid-2000s. This movement of national output is sometimes called the *business cycle*, but it must be recognized that these “cycles” are far from regular or consistent.

Each cycle can be divided by definition into four phases, as shown in Figure B.6. *The trough is the point where national output is lowest relative to its full-employment level. Expansion is the subsequent phase during which national output rises. The peak occurs when national output is highest relative to its full-employment level. Finally, recession is the subsequent phase during which national output falls.*⁶

Many business and economic time series go up and down with the business cycle. For example, industrial output tends to be above its trend line at the peak of the business cycle and tends to fall below its trend line at the trough. Similarly, such diverse series as the money supply, industrial employment, and stock prices reflect the business cycle. However, not all series go up and down at exactly the same time. Some turn upward before others at a trough, and some turn downward before others at a peak. As we shall see, the fact that some time series tend to precede others in cyclical variation sometimes is used to forecast the pace of economic activity.

The **business cycle** or **business fluctuations** reflect the periodic ups and downs experienced by an economy. It can be measured with any economic variable but the fluctuation of GDP over time is the most common.

The **full-employment level** of GDP is the total amount of goods and services that could have been produced if there had been full employment.

The **trough** of the business cycle is where GDP is lowest relative to full-employment GDP.

The **expansion** phase of the business cycle is when GDP rises (such as after reaching the trough).

The **peak** of the business cycle is where GDP is highest relative to full employment GDP.

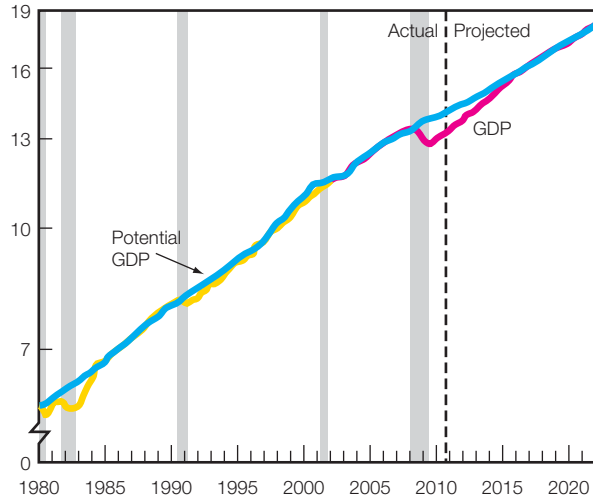
The **recession** phase of the business cycle is when GDP falls (such as after reaching the peak).

5. During a period of inflationary pressure, national output may exceed its full-employment level.

6. The peak and trough may also be defined in terms of deviations from the long-term trend of GDP rather than in terms of deviations from the full-employment level of GDP.

FIGURE B.5

Gross Domestic Product (GDP) (1987 Dollars), United States, 1980–2010

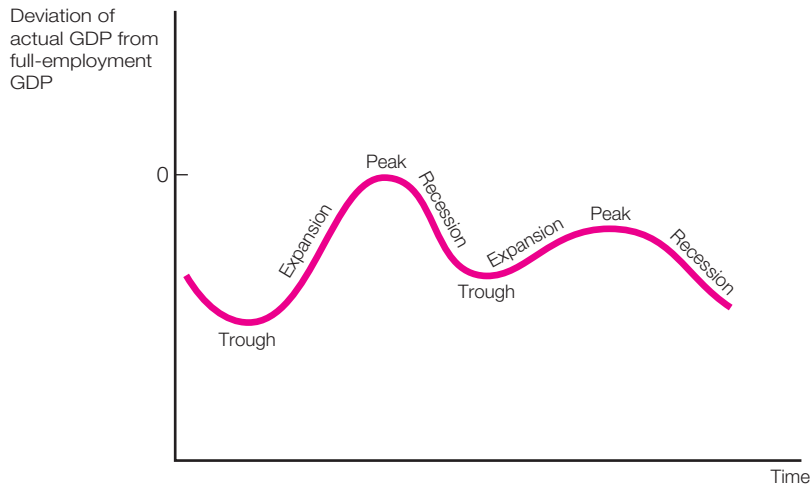


Source: Congress of the United States, Congressional Budget Office, The Budget and Economic Outlook: Fiscal Year 2011 to 2021, January 2011, p. 28.

FIGURE B.6

Four Phases of Business Fluctuations

The peak occurs when national output is highest relative to its full-employment value; the trough occurs when national output is lowest relative to its full-employment value.



ELEMENTARY FORECASTING TECHNIQUES

In general, all forecasting techniques are extremely fallible, and all forecasts should be treated with caution. Nonetheless, businesses and government agencies have no choice but to make forecasts, however crude. Since firms, governments, and private individuals must continually make decisions that hinge on what they expect will happen, they must make implicit forecasts even if they do not make explicit ones. Therefore, the central question is how best to forecast, not whether to forecast. In this section, we present some elementary forecasting techniques that are commonly applied. Even among small firms, evidence indicates that about three-quarters of the firms use techniques of this sort. However, these techniques should be viewed as crude first approximations rather than highly sophisticated methods. More sophisticated techniques are taken up subsequently.

The simplest type of forecasting method is a straightforward extrapolation of a trend. For example, let us return to the ABC Corporation. At the end of 2011, suppose that ABC Corporation managers wanted to forecast its 2012 sales. During the period 1997–2011, we know from our earlier discussion that the firm's sales could be represented (approximately) by the trend line

$$S_t = -11,812 + 5.966t$$

where t equals the year in question. To forecast its 2012 sales, the ABC Corporation could simply insert 2012 for t in this equation. Thus, the forecast for 2012 is

$$-11,812 + 5.966(2012) = 191.6$$

or 191.6 billion dollars. As shown in Figure B.7, this forecast is a simple extension, or extrapolation, of the trend line into the future.

Decision makers often need forecasts of monthly rather than annual amounts. In such cases, it is necessary to recognize that seasonal variation, as well as trend, is likely to affect the value for a particular month. To see how a forecast can be made under such circumstances, consider a clothing manufacturer that wants to forecast its sales during each month of 2012. On the basis of data for each month during the period 1981 to 2011, the firm determines that its sales seem to conform to the trend

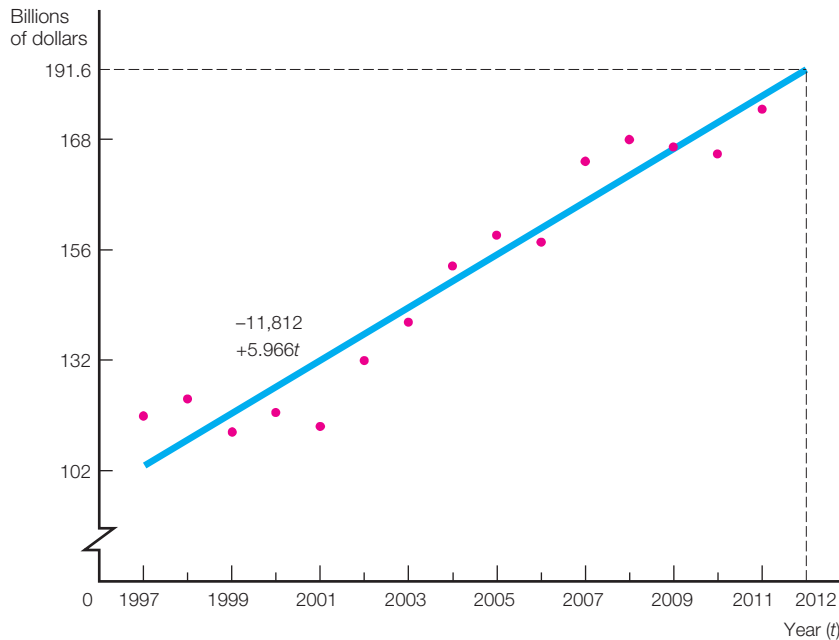
$$S_t = 12,030 + 41t$$

where S_t is the trend value of the firm's monthly sales (in thousands of dollars) and t is time measured in months from January 2011. If this trend continues, the expected sales for each month in 2012 would be as shown in the second column of Table B.2. But this ignores whatever seasonal variation exists in the firm's sales. To include seasonal variation, suppose that the clothing manufacturer's marketing manager analyzes past sales data and finds that the monthly seasonal index for

FIGURE B.7

Simple Trend Extrapolation to Forecast 2012 Sales of the ABC Corporation

The forecast is \$191.6 billion.



sales is as shown in the third column of Table B.2. (Seasonal effects here are multiplicative, not additive.) If this seasonal pattern continues in 2012 as in the past, we would expect that actual sales each month would equal the trend value (in the second column) times the seasonal index (in the third column) divided by 100. The result, which is shown in the fourth column of Table B.2, is a forecast that includes both the trend and the seasonal variation.

Needless to say, this entire procedure is simply a mechanical extrapolation of the firm's sales data into the future. The assumption is made that the past trend and the past seasonal variation will continue. Moreover, it is assumed that the trend and seasonal variation are the predominant factors that will determine sales in the coming months. The validity of this assumption depends on many considerations, including the extent to which the time series in question (in this case, sales) is affected by cyclical factors and the extent to which the economy is likely to change its cyclical position. In the next section, we turn our attention to a particular method of forecasting business fluctuations.

HOW LEADING INDICATORS ARE USED

Managers and analysts want to modify their forecasts to reflect prospective overall changes in economic activity. For example, if the president of the clothing firm in Table B.2 is convinced that a serious depression will occur 2012, he is likely to modify the forecasts in Table B.2 accordingly. But how does the president of the clothing firm—or anyone else—predict whether there is going to be a depression? There are a variety of ways of doing this, all of which are very imperfect. In this section, we discuss an essentially empirical approach, reserving a discussion of more sophisticated techniques for a later section.

One of the simplest ways to forecast business fluctuations is to use **leading indicators**, which are certain economic series that typically go down or up before gross domestic product does. The National Bureau of Economic Research carries out detailed and painstaking examinations of the behavior of various economic variables over a long period of time and attempts to find out whether each variable turns downward before, at, or after the peak of the business cycle and whether it turns upward before, at, or after the trough. Variables that go down before the peak and up before the trough are called **leading series**. Variables that go down at the peak and up at the trough are called **coincident series**. Variables that go down after the peak and up after the trough are called **lagging series**.

Leading indicators are certain economic series that typically go up or down before GDP does.

A **leading series** are variables that go down before the peak and up before the trough.

A **coincident series** are variables that go down at the peak and up at the trough.

A **lagging series** are variables that go down after the peak and up after the trough.

TABLE B.2

Forecast of Sales of Clothing Manufacturer, 2012

Month	Forecast trend value of sales ^a	Seasonal index	Forecast sales (reflecting both trend and seasonal variables) ^a
January	12,522	90	11,270
February	12,563	80	10,050
March	12,604	80	10,083
April	12,645	90	11,380
May	12,686	110	13,955
June	12,727	120	15,272
July	12,768	80	10,214
August	12,809	110	14,090
September	12,850	120	15,420
October	12,891	100	12,891
November	12,932	100	12,932
December	12,973	120	15,568

^aExpressed in units of \$1,000.

According to the bureau, some important leading series are new orders for durable goods, the average workweek, building contracts, stock prices, certain wholesale prices, and claims for unemployment insurance. These are the variables that tend to turn downward before the peak and upward before the trough.⁷ Coincident series include employment, industrial production, corporate profits, and gross domestic product, among many others. Some typical lagging series are retail sales, manufacturers' inventories, and personal income.

These leading series—or leading indicators, as they often are called—are used frequently as forecasting devices. There are sound economic reasons why these series turn downward before a peak or upward before a trough: In some cases, leading series indicate changes in spending in strategic areas of the economy, while in others, they indicate changes in managers' and investors' expectations. To guide business executives in their planning, it is important to try to spot turning points—peaks and troughs—in advance. This, of course, is the toughest part of economic forecasting. Economists sometimes use leading indicators as evidence that a turning point is about to occur. If a large number of leading indicators turn down, this is viewed as a sign of a coming peak. The upturn of a large number of leading indicators is thought to signal an impending trough.

Experience with leading indicators has been only partially successful. The economy has seldom turned downward in recent years without a warning from these indicators, but unfortunately these indicators have turned down on a number of occasions—in 1952 and 1962, for example—when the economy did not turn down subsequently. Therefore, leading indicators sometimes provide false signals. Also, in periods of expansion, they sometimes turn downward too far ahead of the real peak. And in periods of recession, they sometimes turn upward only a very short while before the trough, so that we turn the corners before anything can be done. Nonetheless, leading indicators are not worthless; they are watched closely and used to supplement other, more sophisticated forecasting techniques.

Students interested in learning more about these leading indicators should visit the website of The Conference Board (www.conference-board.org). The Conference Board is a nonprofit organization that provides economic reports and data about business issues and sponsors the business cycle indicators website. The website not only shows a time series for the leading indicators but also discusses their implications for the U.S. economy. In addition, the site reports the leading indicators for other countries.

HOW ECONOMETRIC MODELS ARE USED

Managers and analysts have tended in recent years to base their forecasts more and more on multiple regression techniques and multiequation models. Increased emphasis has been put on the construction and estimation of an equation or system of equations to show the effects of various independent variables on the

7. Of course, claims for unemployment insurance turn upward before the peak and downward before the trough.

variable or variables one wants to forecast. For example, one may want to estimate the quantity of automobiles produced by U.S. auto firms next quarter. According to a study published by the Federal Reserve Bank of New York, the following regression equation is useful for this purpose

$$A = -22,302 + 12.9D - 97.8I - 19.9R + 230P + 6.0N$$

where A is the quantity of autos produced quarterly, D is real disposable income, I is the prime interest rate, R is the inventory-sales ratio, P is the auto price, and N is the nonauto price level. To forecast the quantity of autos produced quarterly, one estimates the values of the independent variables and inserts them into this equation.

PROBLEM SOLVED: Forecasting Shipments of Cement by CEMCO

CEMCO, a small cement producer, has used an econometric model to forecast its sales and profits. According to this model, national cement shipments depend on the amount of residential construction and business fixed investment. Assuming its price is unchanged, CEMCO's shipments of cement are assumed to depend on national cement shipments. Holding national cement shipments constant, CEMCO can increase its shipments by reducing its price. However, its rivals are likely to meet such a price cut, whereas they are less likely to match a price increase.

This year, CEMCO shipped 453,000 tons of cement. Based on this model and alternative assumptions concerning the firm's future price, the forecasted shipments for next year and the year after next were as follows (in thousands of tons):

Assumed future change in CEMCO's price	Next year	Year after next
No price change	468	457
10% price increase	306	296
10% price decrease	473	459

(a) As stated, the firm's model assumes that (1) national cement shipments depend on the amount

of residential construction and business fixed investment, and (2) its shipments depend on national cement shipments (if its price does not change). Does it appear that the company expected both residential construction and business fixed investment to be higher in the year after next than in next year? Why or why not? (b) With regard to increases in price, does the demand for the firm's cement seem to be price elastic or price inelastic? Explain. (c) Does the price elasticity of demand seem to be lower for price decreases than for price increases? Is this reasonable? Why or why not?

SOLUTION (a) No. If the company had expected both residential construction and business fixed investment to be higher in the year after next than in next year, it would have forecasted an increase in national cement shipments, which in turn would have implied an increase in CEMCO's cement shipments (assuming no price change). In fact, as shown in the table, it forecasted that its cement shipments would be lower in the year after next than in next year. (b) Price elastic. A 10% increase in price seems to reduce shipments by about one-third. (c) Yes. It seems reasonable if, as stated, the firm's rivals are likely to meet a price reduction but unlikely to meet a price increase.

Multiequation models have been used to forecast many variables, such as gross domestic product. The Wharton model, a pioneer in this field, contained hundreds of equations variously intended to explain the level of expenditures by households, the level of business investment, aggregate output and employment, and wages, prices, and interest rates. The forecasts produced by the Wharton model (and other large models like it) have been followed closely by major business firms and government agencies. Indeed, some firms (like General Electric) have constructed their own multiequation models. Of course, this does not mean that these large models have an unblemished forecasting record; on the contrary, they, like all other forecasting techniques, are quite fallible. However, these models continue to be used in business and government.

Both the single-equation model used to forecast the quantity of autos produced and the Wharton model, with its hundreds of equations, are examples of econometric models. An **econometric model** is a system of equations (or a single equation) estimated from past data used to forecast economic and business variables. The essence of any econometric model is that it blends economic theory with modern statistical methods.

An **econometric model** is a system of equations estimated from past data used to forecast economic and business variables.

THE PURVERE CORPORATION: A NUMERICAL EXAMPLE

To illustrate the nature of multiequation econometric models, consider the Purvere Corporation, a seller of aircraft. Purvere's total revenues come from three sources: the sale of the equipment, servicing the equipment, and the sale of accessories to customers who buy equipment or have it serviced. On the basis of regression analysis, Purvere's managers have found that its revenues from each of these sources can be represented by the following three equations

$$E_t = 100 - 4P_t + 0.02G_t \quad (\text{B.10})$$

$$S_t = 10 + 0.05E_{t-1} \quad (\text{B.11})$$

$$A_t = 25 + 0.1Y_t \quad (\text{B.12})$$

where E_t is the company's revenue from equipment sales in year t , P_t is the price of its equipment, G_t is gross domestic product (in billions of dollars), S_t is its revenue from servicing its equipment, A_t is its revenue from accessory sales, and Y_t is its total sales (which equal $E_t + S_t + A_t$). The values of E_t , S_t , A_t , and P_t are expressed in millions of dollars.

According to equation (B.10), Purvere's equipment sales are inversely related to its price and directly related to gross domestic product. According to equation (B.11), its service revenues are directly related to its equipment sales during the previous year (because the equipment is serviced about a year after it is bought). According to equation (B.12), its revenue from accessory sales is directly related to its total sales.

Purvere's president wants to use this model to forecast next year's total sales, which equal (in year t)

$$Y_t = E_t + S_t + A_t = (100 + 10 + 25) - 4P_t + 0.02G_t + 0.05E_{t-1} + 0.1Y_t$$

Therefore

$$(1 - 0.1)Y_t = 135 - 4P_t + 0.02G_t + 0.05E_{t-1}$$

or

$$Y_t = \frac{1}{0.9}(135 - 4P_t + 0.02G_t + 0.05E_{t-1}) \quad (\text{B.13})$$

This equation can be used to forecast next year's value of Y if we know the price of Purvere's equipment next year, the value of GDP next year, and the firm's revenues from equipment sales this year. Suppose that price will be 10 and that this year's equipment sales will be 100. Then

$$\begin{aligned} Y_t &= \frac{1}{0.9}(135 - 4(10) + 0.05(100) + 0.02G_t) \\ &= \frac{1}{0.9}(100 + 0.02G_t) \end{aligned}$$

To forecast Y_t , we must know G_t , the value of gross domestic product next year. Obviously, the best we can do is to utilize the best available forecast of next year's GDP. Suppose that Purvere's president decides to rely on the forecast of a large econometric model (like Wharton's), which is that GDP next year will be about \$6,250 billion. If so, his sales forecast for next year would be

$$Y_t = \frac{1}{0.9}[100 + 0.02(6,250)] = \frac{1}{0.9}(225) = 250$$

or \$250 million.

Note that Purvere's president links his company model in equations (B.10) to (B.12) to the large econometric model, which is providing the forecasted value of G_t . This is frequently the way managers have used the forecasts of macroeconomic models like the Wharton model.

Before leaving this example, it is important to recognize that it is highly simplified. Firms frequently use multiequation models containing many variables, not just the handful contained in equations (B.10) to (B.12).

“STUDY YOUR RESIDUALS”

Before concluding this appendix, it is important to consider Nobel laureate Paul A. Samuelson's well-known statement: “To the scientific forecaster I say, ‘Always study your residuals.’” What Samuelson meant was that, in evaluating any fore-

casting technique, it is useful to calculate the difference between each observation and what the technique predicts this observation will be. These differences—or residuals—are very useful in indicating whether your forecasting technique excludes some important explanatory variables and whether its assumptions are valid.

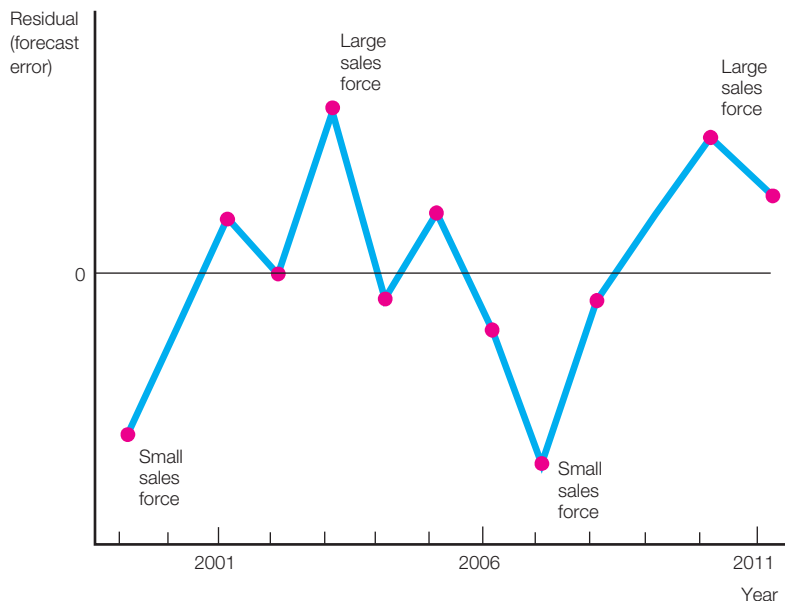
To illustrate, suppose you are using an econometric model to forecast your firm’s sales and the difference between each year’s sales and what this model predicts these sales to be is as shown in Figure B.8. To improve this technique, you should think hard about why the model made the errors that it did. Based on Figure B.8, it might occur to you, for example, that many of the years when the residuals were large and positive were years when your firm had an unusually large sales force and that many of the years when the residuals were large and negative were years when your firm had an unusually small sales force. If the size of your firm’s sales force is not included as an independent variable in your model, there may be good reason to include it.

By continually studying your forecasting errors and improving your forecasting techniques, significant progress can be made. Although it generally is unrealis-

FIGURE B.8

Residuals from Sales Forecasting Model

The years (1999 and 2007) when the residuals are large and negative are ones when the sales force was small; the years (2003 and 2010) when they are large and positive are ones when the sales force was large.



tic to expect business and economic forecasts to be very precise, they are likely to be considerably more trustworthy than forecasts that are not based on the principles of managerial economics.

SUMMARY

1. Although surveys are of considerable use, most major firms seem to base their forecasts in large part on the quantitative analysis of economic time series. The classical approach to business forecasting assumes that an economic time series can be decomposed into four components: trend, seasonal variation, cyclical variation, and irregular movements.

2. If the trend in a time series is linear, simple regression may be used to estimate an equation representing the trend. If it seems to be nonlinear, a quadratic equation may be estimated by multiple regression or an exponential trend may be fitted. An exponential trend is appropriate when the variable increases at a relatively constant percentage rate per year. To fit such a trend, we use the logarithm of the variable, not the variable itself, as the dependent variable in the regression.

3. The seasonal variation in a particular time series is described by a figure for each month (the seasonal index) that shows the extent to which that month's value typically departs from what would be expected on the basis of trend and cyclical variation. Such seasonal indexes, together with the trend, can be useful for forecasting. Regression analysis, including dummy variables, can be employed to estimate seasonal indexes.

4. Cyclical variation, as well as trend and seasonal variation, is reflected in many time series. Variables that go down before the peak and up before the trough are called *leading indicators*. If a large number of leading indicators turn downward, this is viewed as a sign of a coming peak. If a large number turn upward, this is thought to signal an impending trough. Although these indicators are not very reliable, they are watched closely and are used to supplement other, more-sophisticated forecasting techniques.

5. The simplest kind of forecasting method is a straightforward extrapolation of a trend. To allow for seasonal variation, either multiplicative or additive seasonal effects can be included. This entire procedure is simply a mechanical extrapolation of the time series into the future.

6. In recent years, managerial economists have tended to base their forecasts less on simple extrapolations and more on equations (or systems of equations) showing the effects of various independent variables on the variable (or variables) one wants to forecast. These equations (or systems of equations), are called *econometric models*. Examples are the models used by CEMCO and General Electric, as well as the model of auto output published by the Federal Reserve Bank of New York.

PROBLEMS

1. The following seasonal index was calculated for the room occupancy of a motel located on a major interstate highway in the Southeast. The motel's customers are largely tourists and commercial truckers who regularly travel this highway. This index is based on actual data.⁸

January	74.8	July	116.8
February	79.8	August	117.4
March	92.9	September	105.4
April	108.8	October	103.7
May	107.5	November	100.3
June	112.0	December	80.6

- Is there pronounced seasonal variation in this motel's business? All other things equal, by what percentage, on the average, does room occupancy in the peak month exceed that in the lowest month?
 - What factors would you expect to be responsible for the observed seasonal variation? In calculating the seasonal index, it was assumed that the index for a particular month like January was the same from year to year. Some observers have questioned whether this assumption is correct, given that a recession occurred in 2008 and 2009. Why might the recession have changed the pattern of seasonal variation?
 - If you were the motel's manager, how might a seasonal index of this sort be of use? Be specific.
2. The Carbide Corporation's sales during the period 1997 to 2012 follow:

Year	Sales (billions of dollars)	Year	Sales (billions of dollars)
1997	1.5	2005	2.7
1998	1.6	2006	2.9
1999	1.6	2007	3.0
2000	1.7	2008	3.0
2001	1.9	2009	3.3
2002	2.1	2010	3.9
2003	2.2	2011	5.3
2004	2.5	2012	5.7

8. B. Bettegowda, "Calculation of Seasonal Index for Motel Room Occupancy," National Technological University, 1991, found at drgeorgejohnny.com/multimedia/lecture/.../07/.doc.

- a. Fit a linear trend line to these data.
 - b. Fit an exponential trend line to these data.
 - c. Assume that in 2020, Carbide's sales are \$9.994 billion. Suppose that in 2012, both the linear trend line and the exponential trend line had been used to forecast the firm's 2020 sales. Which forecast would have been more accurate?
 - d. Assume that in 2025, Carbide's sales are \$9.508 billion. Suppose that in 2012, both the linear trend line and the exponential trend line had been used to forecast the firm's 2025 sales. Which forecast would have been more accurate?
3. The Milton Company's statistician calculates a seasonal index for the firm's sales; the results are shown in the second column. The firm's monthly 2011 sales are shown in the third column.

Month	Seasonal index	2011 sales (millions of dollars)
January	97	2.5
February	96	2.4
March	97	2.7
April	98	2.9
May	99	3.0
June	100	3.1
July	101	3.2
August	103	3.1
September	103	3.2
October	103	3.1
November	102	3.0
December	101	2.9

- a. If one divides each month's sales figure by its seasonal index (divided by 100), it is said to be "deseasonalized." That is, the seasonal element is removed from the data. Why is this true?
 - b. Calculate deseasonalized sales figures for 2011.
 - c. Why would the managers of the Milton Company want deseasonalized sales figures?
4. The equation describing the sales trend of the Secane Chemical Company is

$$S_t = 21.3 + 1.3t$$

where S_t is the sales (in millions of dollars per month) of the firm and t is time measured in months from January 2012. The firm's seasonal index of sales is

January	103	May	101	September	121
February	80	June	104	October	101
March	75	July	120	November	75
April	103	August	139	December	78

- a. Construct a monthly sales forecast for the firm for 2013.
 - b. Why would the managers of the Secane Chemical Company want monthly sales forecasts of this kind?
5. On October 4, 2011, the U.S. Department of Commerce announced that the index of leading indicators rose 0.6% in August 2011.
- a. During August, the average workweek rose. Is the average workweek among the leading indicators? If so, did its increase help to raise the index?
 - b. During August, stock prices rose. Is the level of stock prices among the leading indicators? If so, did its increase help to raise the index?
6. The Allen Company's monthly sales have the following trend

$$C_t = 4.12 + 0.32t$$

where C_t is the sales (in millions of dollars per month) of the firm and t is time measured in months from July 2004. The firm's seasonal index of sales is

January	81	May	137	September	79
February	98	June	122	October	101
March	102	July	104	November	74
April	76	August	101	December	125

- a. Construct a monthly sales forecast for the firm in 2012.
 - b. The firm's president feels strongly that a recession will occur in late 2012. Would this influence your answer to part a? If so, how?
7. The sales of Sears, Roebuck were as follows during 1978 to 1990:

Year	Sales (billions of dollars)	Year	Sales (billions of dollars)
1978	22.9	1985	40.7
1979	24.5	1986	42.3
1980	25.2	1987	45.9
1981	27.4	1988	50.3
1982	30.0	1989	53.8
1983	35.9	1990	56.0
1984	38.8		

- a. Calculate a linear trend based on these data.
 - b. Sears, Roebuck's sales in 1991 were about \$57.2 billion. If you had used a least-squares trend line based on 1978 to 1990 data to forecast its 1991 sales, how big a forecasting error would have resulted?
 - c. In 1992, Sears, Roebuck's sales were \$52.3 billion. If you had used a least-squares trend line based on 1978 to 1990 data to forecast its 1992 sales, how big a forecasting error would have resulted?
8. In the Wharton econometric model, housing starts (divided by the number of households) were specified to be a function of (1) the mortgage rate, (2) a consumer sentiment index, (3) capacity utilization, (4) the occupancy rate, and (5) deposit inflows into savings intermediaries.
- a. Indicate why each of these five variables might be expected to influence the number of housing starts.
 - b. What factors influence these five variables? What sort of multiequation system might be constructed for forecasting purposes?
9. A firm's sales from 1986 to 2012 were as follows:

Year	Sales (billions of dollars)	Year	Sales (billions of dollars)
1986	2.2	2000	4.9
1987	2.6	2001	6.2
1988	3.0	2002	7.2
1989	3.5	2003	7.7
1990	3.3	2004	8.4
1991	3.5	2005	8.4
1992	4.1	2006	8.8
1993	4.3	2007	9.6
1994	4.2	2008	10.5
1995	4.5	2009	11.9
1996	4.2	2010	13.9
1997	4.5	2011	14.1
1998	4.8	2012	15.7
1999	4.9		

- a. Using the method of least squares, derive a linear trend.
- b. Plot the firm's sales against time. Also, plot the trend line derived in part a against time. (Time here is the year to which the sales figure pertains.)
- c. Does a visual inspection of how well the linear trend fits suggest that an exponential trend would do better? That a quadratic trend would do better?
- d. Using this linear trend, what would have been the sales forecast for the firm in 2015?

APPENDIX: EXPONENTIAL SMOOTHING AND FORECASTING

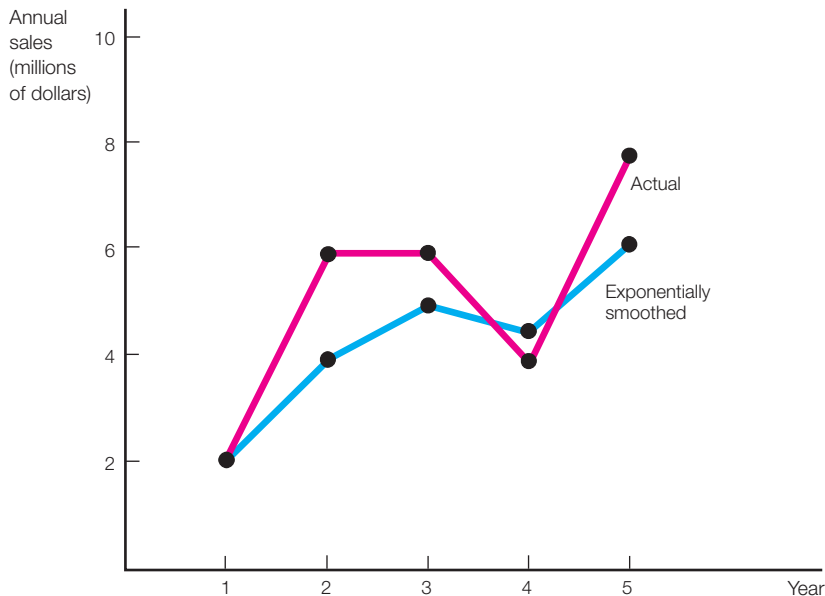
A frequently used method of calculating a trend is by **exponential smoothing**. According to this method, *the trend value at time t is a weighted average of all available previous values, where the weights decline geometrically as one goes backward in time*. As an illustration, suppose that a firm has been in existence for five years and its sales have been \$2 million, \$6 million, \$6 million, \$4 million, and \$8 million (see Figure B.9). Then, the trend value in the fifth year would be a weighted average of \$2 million, \$6 million, \$6 million, \$4 million, and \$8 million, where the weights decline geometrically as we go backward in time. Specifically, the weight attached to the observation at time t equals θ , the weight attached to the observation at time $t - 1$ equals $(1 - \theta)\theta$, the weight attached to the observation at time $t - 2$ equals $(1 - \theta)^2\theta$, the weight attached to the observation at time $t - 3$ equals $(1 - \theta)^3\theta$, ..., and the weight attached to the observation at the earliest relevant time (time 0) equals $(1 - \theta)^t$. Clearly, the weights decline geometrically as one goes backward in time; that is, the weight attached to the observation at time $t - 1$ is $(1 - \theta)$ times the weight attached to the observation at time t ; the weight attached to the observation at time $t - 2$ is $(1 - \theta)$ times the weight attached to the observation at time $t - 1$; and so on.

Exponential smoothing occurs when the trend value at time t is a weighted average of all available previous values (where the weights decline geometrically as one goes backward in time).

FIGURE B.9

Sales of Firm, Actual and Exponentially Smoothed

Year 1 is the firm's first year in existence, year 2 is its second year, and so on.



The **smoothing constant** is an arbitrary value [between 0 and 1] chosen by a manager which determines the weight a manager assigns to an observation at time t . This, in turn, determines the weight assigned to each observation in previous periods.

To calculate an exponentially smoothed time series, choose a value of θ , which is designated the **smoothing constant**. If we choose a value of 0.5 for θ , the exponentially smoothed value of the firm's sales in each of the five years is

$$\begin{aligned} S_0 &= 2 \\ S_1 &= (0.5)(6) + (1 - 0.5)(2) = 4 \\ S_2 &= (0.5)(6) + (1 - 0.5)(0.5)(6) + (1 - 0.5)^2(2) = 5 \\ S_3 &= (0.5)(4) + (1 - 0.5)(0.5)(6) \\ &\quad + (1 - 0.5)^2(0.5)(6) + (1 - 0.5)^3(2) = 4.5 \\ S_4 &= (0.5)(8) + (1 - 0.5)(0.5)(4) + (1 - 0.5)^2(0.5)(6) \\ &\quad + (1 - 0.5)^3(0.5)(6) + (1 - 0.5)^4(2) = 6.25 \end{aligned}$$

where S_0 is the exponentially smoothed value of the firm's sales in the first year of its existence, S_1 is this value in the second year, S_2 the value in the third year, and so on. Figure B.9 shows both the original time series and the exponentially smoothed time series.

To compute the value of such a smoothed time series at time t , all you really need is the value of the smoothed time series at time $t - 1$ and the actual value of the time series at time t . This is because the smoothed value of the time series at time t is a simple weighted average of the smoothed value at time $t - 1$ and the actual value at time t . If S_t is the smoothed value at time t

$$S_t = \theta Y(t) + (1 - \theta)S_{t-1} \quad (\text{B.14})$$

where $Y(t)$ is the value of the time series at time t .⁹ So, to calculate an exponentially smoothed time series, you do not need to keep all the previous values of the actual time series; *all you need to keep is the value of the exponentially smoothed series in the previous period*. From this information alone (together with the current value of the series and the smoothing constant), you can calculate the smoothed value of the series in the current period. For instance, consider the firm in the previous paragraph. If the firm's sales in its sixth year of existence are \$10 million, the smoothed value of sales for the sixth year is

$$(0.5)(10) + (1 - 0.5)(6.25) = 8.125$$

or \$8.125 million.

In choosing the value of the smoothing constant θ , you must pick a number between 0 and 1. (In other words, $0 \leq \theta \leq 1$.) If θ is close to 1, past values of the time series are given relatively little weight (compared with recent values) in calculating smoothed values. If θ is close to 0, past values of the time series are given considerable weight (as compared with recent values) in calculating smoothed values. If the time series contains a great deal of random variation, it is often advisable to choose a relatively small value of θ , since this results in relatively little weight put on $Y(t)$, which is influenced more than S_{t-1} by this variation. On the other hand, if one wants the smoothed time series to reflect relatively quickly

9. Let us prove that equation (B.14) is true. If $Y(t)$ is the actual value of the time series at time t , then equation (B.14) implies that

$$\begin{aligned} S_t &= \theta Y(t) + (1 - \theta)S_{t-1} \\ &= \theta Y(t) + (1 - \theta)[\theta Y(t-1) \\ &\quad + (1 - \theta)S_{t-2}] \\ &= \theta Y(t) + (1 - \theta)\theta Y(t-1) \\ &\quad + (1 - \theta)^2[\theta Y(t-2) \\ &\quad + (1 - \theta)S_{t-3}] \\ &\vdots \\ &= \theta Y(t) + (1 - \theta)\theta Y(t-1) \\ &\quad + (1 - \theta)^2\theta Y(t-2) \\ &\quad + \dots + (1 - \theta)^{t-1}Y(0) \end{aligned}$$

Since the right-hand side of the last line is equivalent to the definition of an exponentially smoothed time series in the first paragraph of this appendix, it follows that equation (B.14) is true.

whatever changes occur in the average level of the time series, the value of θ should be set at a high level.

Forecasting Based on Exponential Smoothing

Exponential smoothing is also used for forecasting purposes. When used in this way, the basic equation for exponential smoothing is

$$F_t = \theta A(t-1) + (1 - \theta)F_{t-1} \quad (\text{B.15})$$

where $A(t-1)$ is the actual value of the time series at time $(t-1)$ and F_t is the forecast for time t . Because the forecast is being made at time $(t-1)$, the actual value of the time series at this time is known. The forecast for time t is simply a weighted average of the actual value at time $(t-1)$ and the forecasted value for time $(t-1)$, where the actual value is weighted by θ and the forecasted value is weighted by $(1 - \theta)$. It can readily be demonstrated that the forecast for time t is the weighted sum of the actual values prior to time t , where the weight attached to each value declines geometrically with the age of the observation.

To see how exponential smoothing can be used for forecasting purposes, we return to the firm in Figure B.9, which had been in existence for five years. Sales during the first year were \$2 million, and we assume that the firm's sales forecast for the first year was also \$2 million. What will be its sales forecast for the second year? To make such a forecast, the firm begins by choosing a value for the smoothing constant θ . (Values of 0.3 or less are often used.) Suppose that a value of 0.2 is chosen. Then the forecast for the second year is $0.2(2) + 0.8(2) = 2$, or \$2 million. Since the firm's actual sales in the second year turn out to be \$6 million, its sales forecast for the third year will be $0.2(6) + 0.8(2) = 2.8$, or \$2.8 million. Since the firm's actual sales in the third year turn out to be \$6 million, its sales forecast for the fourth year will be $0.2(6) + 0.8(2.8) = 3.44$, or \$3.44 million. And so on. Exponential smoothing is often used in this way to make forecasts, particularly where there is a need for a cheap, fast, and rather mechanical method to make forecasts for a large number of items. For example, to implement various kinds of inventory control models, demand forecasts for hundreds or thousands of items may be required.

LEARNING OBJECTIVES

Present Value of a Series of Payments

The Use of Periods Other Than a Year

Determining the Internal Rate of Return

DISCOUNTING AND PRESENT VALUES

When a manager chooses between two courses of action, *A* and *B*, he or she is choosing between the cash flows resulting if *A* is chosen and the cash flows if *B* is chosen. These cash flows generally occur over a number of periods. For example, if *A* is chosen, the manager may experience an outflow of \$1 million this year and an inflow of \$300,000 during each of the next five years. On the other hand, if *B* is chosen, the manager may experience an outflow of \$1 million this year and an inflow of \$250,000 for each of the next six years. How can a manager compare these two alternatives?

To answer this question, it is convenient to begin by pointing out one of the basic propositions in managerial economics: *A dollar received today is worth more than a dollar received a year from today.* Why? Because one can always invest money that is available now and obtain interest on it. If the interest rate is 6%, a dollar received now is equivalent to \$1.06 received a year hence. Why? Because if you invest the dollar now, you'll get \$1.06 in a year. Similarly, *a dollar received now is equivalent to (1.06)² dollars two years hence.* Why? Because if you invest the dollar now, you'll get 1.06 dollars in a year, and if you reinvest this amount for another year at 6%, you'll get (1.06)² dollars.

More generally, suppose you can invest at a compound rate of *i*% per year. What is the *present value*—that is, the value *today*—of a dollar received *n* years hence? Based on the foregoing argument, its present value is

$$\frac{1}{(1 + i)^n} \quad (\text{C.1})$$

Thus, if the interest rate is 0.10 and if $n = 4$ (which means that the dollar is received in four years), the present value of a dollar equals

$$\frac{1}{(1 + 0.10)^4} = \frac{1}{1.4641} = \$0.683$$

In other words, the present value of the dollar is 68.3 cents.

To see that this answer is correct, let's see what would happen if you invested 68.3 cents today. As shown in Table C.1, this investment would be worth 75.1 cents after one year, 82.6 cents after two years, 90.9 cents after three years, and 1 dollar after four years. Thus, 68.3 cents is the present value of a dollar received four years hence, because if you invest 68.3 cents today, you will have exactly 1 dollar in four years.

Table E.1 shows the value of $1/(1 + i)^n$, for various values of i and n . For example, according to this table, the present value of a dollar received ten years hence is 46.3 cents if the interest rate is 0.08. To see this, note that the figure in Table E.1 corresponding to $n = 10$ and $i = 0.08$ is 0.46319.

Using this table, you can readily determine the present value of any amount received n years hence, not just 1 dollar. If you receive R_n dollars n years hence, the present value of this amount is

$$\frac{R_n}{(1 + i)^n} \quad (\text{C.2})$$

Thus, to determine the present value of R_n , all that you have to do is multiply R_n by $1/(1 + i)^n$. Since Table E.1 provides us with the value of $1/(1 + i)^n$, this is a simple calculation.

To illustrate, suppose you will receive \$10,000 ten years hence and the interest rate is 0.12. According to equation (C.2), the present value of this amount equals $\$10,000[1/(1 + i)^n]$. Since Table E.1 shows that $1/(1 + i)^n = 0.32197$ when $n = 10$ and $i = 0.12$, the present value of this amount is $\$10,000(0.32197) = \$3,219.70$.

TABLE C.1

Value of 68.3 Cents Invested at 10% Interest

Number of years hence	Return received	Value of investment
1	$68.301(0.10) = 6.830¢$	$68.301 + 6.830 = 75.13¢$
2	$75.131(0.10) = 7.513¢$	$75.131 + 7.513 = 82.64¢$
3	$82.643(0.10) = 8.264¢$	$82.645 + 8.265 = 90.91¢$
4	$90.907(0.10) = 9.091¢$	$90.909 + 9.091 = 100.00¢$

PRESENT VALUE OF A SERIES OF PAYMENTS

As pointed out at the beginning of this appendix, managers generally must consider situations in which cash flows occur at more than a single time. For example, investment in a new machine tool is likely to result in a cash outflow now and a series of cash inflows in the future. To determine the present value of such an investment, it is convenient to begin by considering the simple case in which you receive \$1 per year for n years, the interest rate being i . More specifically, the n receipts of \$1 occur one year from now, two years from now, . . . , and n years from now. The present value of this stream of \$1 receipts is

$$\frac{1}{1+i} + \frac{1}{(1+i)^2} + \cdots + \frac{1}{(1+i)^n} = \sum_{t=1}^n \frac{1}{(1+i)^t} \quad (\text{C.3})$$

For example, the present value of \$1 to be received at the end of each of the next five years, if the interest rate is 0.10, is

$$\begin{aligned} \sum_{t=1}^5 \frac{1}{(1+0.10)^t} &= \frac{1}{(1+0.10)} + \frac{1}{(1+0.10)^2} + \frac{1}{(1+0.10)^3} \\ &+ \frac{1}{(1+0.10)^4} + \frac{1}{(1+0.10)^5} = 0.90909 + 0.82645 \\ &+ 0.75131 + 0.68301 + 0.62092 = \$3.79 \end{aligned} \quad (\text{C.4})$$

To obtain each of the terms on the right in equation (C.4), we use Table E.1. For example, the final term on the right is 0.62092, which is the present value of a dollar received five years hence (if the interest rate is 0.10), according to Table E.1.

Table C.2 shows that \$3.79 is indeed the present value of \$1 to be received at the end of each of the next five years, if the interest rate is 0.10. As you can see, if you invest \$3.79 at 10% interest, you will be able to withdraw \$1 at the end of each year, with nothing left over or lacking. Since analysts frequently must calculate the present value of a dollar received at the end of each of the next n years, the expression in equation (C.3)

$$\sum_{t=1}^n 1/(1+i)^t$$

has been tabled; the results are shown in Table E.2. For example, if you receive \$1 at the end of each of the next ten years, and if the interest rate is 0.06, the present value is \$7.36. To see this, note that the figure in Table E.2 corresponding to $n = 10$ and $i = 0.06$ is 7.3601.

More generally, if you receive R dollars at the end of each of the next n years, and if the interest rate is i , the present value is

$$\sum_{t=1}^n \frac{R}{(1+i)^t} = R \sum_{t=1}^n \frac{1}{(1+i)^t} \quad (\text{C.5})$$

TABLE C.2

Demonstration that \$3.79 (Invested at 10% Interest) Provides Exactly \$1 at the End of Each of the Next Five Years

Number of years hence	Return received	Amount withdrawn	Net Value of investment
1	$\$3.790(0.10) = \0.3790	\$1.00	$\$3.790 + 0.3790 - 1.00 = \3.169
2	$3.169(0.10) = 0.3169$	\$1.00	$3.169 + 0.3169 - 1.00 = 2.486$
3	$2.486(0.10) = 0.2486$	\$1.00	$2.486 + 0.2486 - 1.00 = 1.735$
4	$1.735(0.10) = 0.1735$	\$1.00	$1.735 + 0.1735 - 1.00 = 0.909$
5	$0.909(0.10) = 0.0909$	\$1.00	$0.909 + 0.0909 - 1.00 = 0$

Thus, the present value of \$5,000 to be received at the end of each of the next five years, if the interest rate is 0.08, is $\$5,000(3.9927) = \$19,963.5$, since Table E.2 shows that the value of $\sum_{t=1}^n 1/(1+i)^t = 3.9927$, when $n = 5$ and $i = 0.08$.

Finally, we must consider the case in which there is a series of unequal, not equal, payments. Suppose that a payment is received at the end of each of the next n years, that the amount received at the end of the t th year is R_t , and that the interest rate is i . The present value of this series of unequal payments is

$$\sum_{t=1}^n \frac{R_t}{(1+i)^t} \quad (\text{C.6})$$

Table E.1 can be used to help carry out this computation. For example, suppose that $i = 0.10$, that $n = 3$, and that the amount received at the end of the first year is \$3,000, the amount received at the end of the second year is \$2,000, and the amount received at the end of the third year is \$1,000. Table C.3 shows how to calculate the present value of this series of unequal payments, which in this case equals \$5,131.48.

THE USE OF PERIODS OTHER THAN A YEAR

Thus far, we have assumed the interest or return from an invested amount is paid annually. In other words, we have assumed that a dollar invested at the beginning of a year earns interest of $i\%$ at the end of that year. In many situations, this is not correct. Instead, interest, dividends, or other returns from an investment may be received semiannually, quarterly, monthly, or even daily. Because you earn a return in the next period on the return received in this period, the results differ from those given in previous sections of this appendix.

TABLE C.3

**Present Value of Stream of Unequal Payments,
Where $i = 0.10$ and $n = 3$**

	(1)	(2)	(1) × (2)
Number of years hence	Amount received R_t	$\frac{1}{(1 + 0.10)^t}$	Present value of amount received
1	\$3,000	0.90909	\$2,727.27
2	2,000	0.82645	1,652.89
3	1,000	0.75131	751.31
			Total \$5,131.48

If interest is received *semiannually*, the present value of a dollar received n years hence is

$$\frac{1}{(1 + i/2)^{2n}} \quad (\text{C.7})$$

where i is the annual interest rate. To understand this expression, note that the interest rate for each semiannual period is $i/2$, and that there are $2n$ semiannual periods in n years. Bearing this in mind, this expression can be derived in the same way as expression (C.1).

If interest is received *quarterly*, the present value of a dollar received n years hence is

$$\frac{1}{(1 + i/4)^{4n}} \quad (\text{C.8})$$

where i once again is the annual interest rate. To see why this is true, note that the interest rate for each quarterly period is $i/4$, and that there are $4n$ quarterly periods in n years. Bearing this in mind, this expression can be derived in the same way as expression (C.1).

More generally, suppose that interest is received c times per year. Under these circumstances, the present value of a dollar received n years hence is

$$\frac{1}{(1 + i/c)^{cn}} \quad (\text{C.9})$$

Table E.1 can be used to determine present values under these circumstances. To evaluate expression (C.9), let the interest rate be i/c , and let the number of years be cn ; using these values, Table E.1 gives the correct answer. Thus, the present value of 1 dollar to be received 3 years hence, where the interest rate is 8% paid

quarterly, can be obtained by finding in the table the present value of 1 dollar to be received 12 years hence where the interest rate is 2%. Specifically, the answer is 78.849 cents.

DETERMINING THE INTERNAL RATE OF RETURN

Previous sections of this appendix have been concerned entirely with determining the present value of a stream of cash flows. While this is of great importance in managerial economics, it also is important to calculate the internal rate of return—the interest rate that equates the present value of the cash inflows with the present value of the cash outflows. Put differently, the internal rate of return is the interest rate that makes the present value of a stream of cash flows equal zero. In other words, we want to find i where

$$R_0 + \frac{R_1}{1+i} + \frac{R_2}{(1+i)^2} + \cdots + \frac{R_n}{(1+i)^n} = 0$$

or

$$\sum_{t=0}^n \frac{R_t}{(1+i)^t} = 0 \quad (\text{C.10})$$

To solve equation (C.10) for i , it often is necessary to use trial and error (if you do not have access to a computer or calculator). The first step is to make a rough estimate of the value of i that will satisfy equation (C.10). The second step is to adjust this estimate. If the present value based on the original estimated rate of interest is *positive*, *increase* the value of i . If the present value based on the original estimated rate of return is *negative*, *reduce* the value of i . The third step is to continue to adjust this estimate until you find the value of i that will satisfy equation (C.10).

As an illustration, consider the following stream of cash flows: $R_0 = -\$5,980$, $R_1 = \$3,000$, $R_2 = \$2,000$, and $R_3 = \$2,000$. As a first step, we estimate (roughly) that the internal rate of return is in the neighborhood of 8%. As Table C.4 shows, the present value of this stream of cash flows, given that the interest rate is 8%, is \$100.12, which is positive. Thus, a higher value of i must be tried. We choose 9%. As Table C.4 shows, the present value of this stream of cash flows, given that the interest rate is 9%, is virtually zero. Thus, the internal rate of return is 9%.

If the cash flows (in years other than year 0) are all equal, there is a simpler way to determine the internal rate of return. Under these circumstances, equation (C.10) can be written

$$R_0 + \sum_{t=1}^n \frac{R}{(1+i)^t} = 0$$

TABLE C.4

Determination of the Internal Rate of Return

Year t	Cash flow R_t	$i = 8\%$		$i = 9\%$	
		$\frac{1}{(1+i)^t}$	Present value	$\frac{1}{(1+i)^t}$	Present value
0	−\$5,980	1.00000	−\$5,980	1.00000	−\$5,980
1	\$3,000	0.92593	\$2,777.78	0.91743	\$2,752.29
2	2,000	0.85734	1,714.68	0.84168	1,683.36
3	2,000	0.79383	1,587.66	0.77228	1,544.37
Total			100.12		0.02

where R is the cash flow in years 1 to n . Thus,

$$\sum_{t=1}^n \frac{1}{(1+i)^t} = \frac{-R_0}{R} \quad (\text{C.11})$$

Since we are given the value of $-R_0/R$, we can find the value of i in Table E.1 where the entry in the n th row equals $-R_0/R$. This value of i is the internal rate of return.

To illustrate, suppose that a machine tool costs \$10,000, and that it will result in a cash inflow of \$2,500 for each of the next six years. Since $R_0 = -\$10,000$ and $R = \$2,500$, the value of $-R_0/R$ is 4. Looking in the row of Table E.2 where $n = 6$, we look for the interest rate where the entry in the table is 6. Since the entry is 3.9976 when $i = 13\%$, the internal rate of return is about 13%.

Finally, it is worth pointing out that if an investment yields an infinite series of equal cash flows, the present value of this series is

$$\sum_{t=1}^{\infty} \frac{R}{(1+i)^t} = R \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} = \frac{R}{i} \quad (\text{C.12})$$

For example, if an investment yields a perpetual annual return of \$4,000 per year, and if the interest rate is 8%, the present value of this perpetual stream of returns equals $\$4,000/0.08 = \$50,000$.

ANSWERS TO SELECT END-OF-CHAPTER PROBLEMS

CHAPTER 1

1. Yes.

3. Number of years in the future	Profit (millions of dollars)	$\frac{1}{(1 + i)^t}$	Present value
1	8	0.90909	7.27272
2	10	0.82645	8.26450
3	12	0.75131	9.01572
4	14	0.68301	9.56214
5	15	0.62092	9.31380
6	16	0.56447	9.03152
7	17	0.51316	8.72372
8	15	0.46651	6.99765
9	13	0.42410	5.51330
10	10	0.38554	3.85540
			Total 77.55056

Thus, the answer is \$77.55056 million.

5. a. He will receive $80(50)(\$5) = \$20,000$, from which he must pay \$3,000 for the umbrellas and $3(\$3,000) = \$9,000$ for rent. Thus, his accounting profit equals $\$20,000 - \$3,000 - \$9,000$ or \$8,000.

- b. Since he could earn \$4,000 doing construction work, his economic profit is $\$8,000 - \$4,000 = \$4,000$. (For simplicity, we ignore the fact that he could have earned interest on the money he invested in this business during the summer.)

CHAPTER 2

1. a. If $Q = 20$, $P = 2,000 - 50(20) = 1,000$. Thus, price would have to equal \$1,000.
 b. Since $500 = 2,000 - 50Q$, $Q = 1,500/50 = 30$. Thus, it will sell 30 per month.
 c. Because $Q = (2,000 - P)/50 = 40 - 0.02P$, $dQ/dP = -0.02$. Thus

$$\left(\frac{P}{Q}\right)\left(\frac{\partial Q}{\partial P}\right) = -0.02 \frac{500}{30} = -0.33$$

- d. If $-0.02 \frac{P}{(2,000 - P)/50} = -1$

$$-0.02 \frac{50P}{2,000 - P} = -1$$

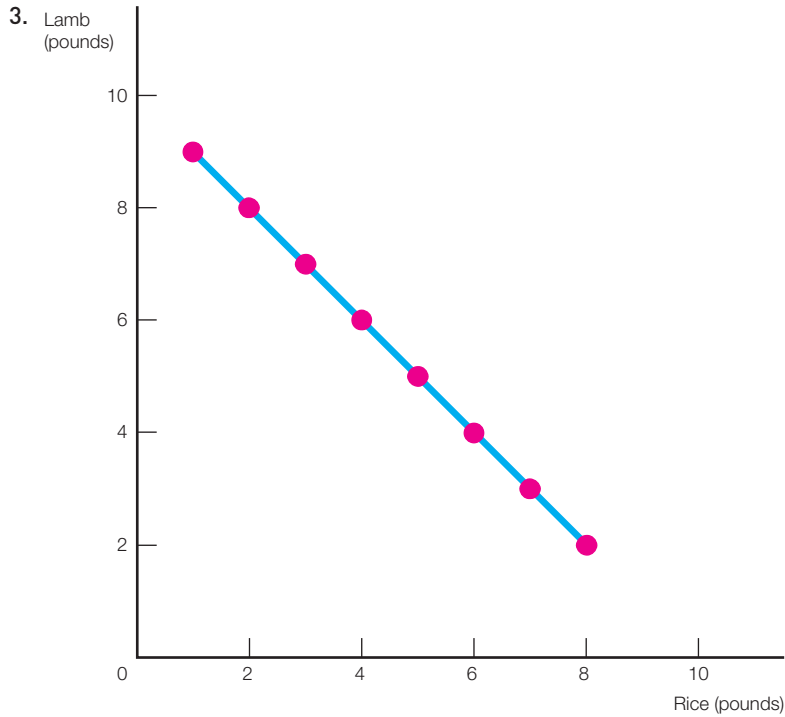
$$\begin{aligned} P &= 2,000 - P \\ &= 2,000/2 = 1,000 \end{aligned}$$

Thus, if price equals \$1,000, the demand is of unitary elasticity.

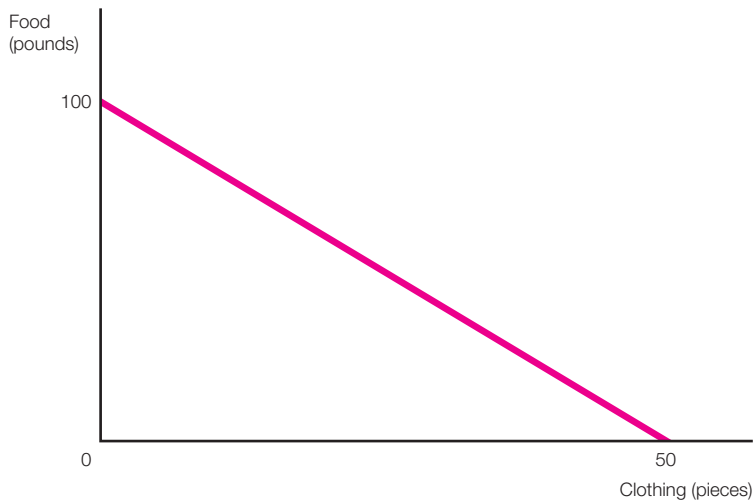
3. a. $\frac{\partial Q}{\partial P} \left(\frac{P}{Q}\right) = \frac{-3(10)}{500 - 3(10) + 2(20) + 0.1(6,000)}$
 $= \frac{-30}{500 - 30 + 40 + 600} = \frac{-30}{1,110}$
 b. $\frac{\partial Q}{\partial I} \left(\frac{I}{Q}\right) = \frac{0.1(6,000)}{1,110} = \frac{600}{1,110}$
 c. $\frac{\partial Q}{\partial P_r} \left(\frac{P_r}{Q}\right) = \frac{2(20)}{1,110} = \frac{40}{1,110}$
 d. Population is assumed to be essentially constant (or to have no significant effect on Q , other than via whatever effect it has on per capita disposable income).
 6. a. Because there are lots of very close substitutes for a particular brand, but not for cigarettes as a whole. It appears that the elasticity was less than -2 .
 b. No. More will be said about estimating demand functions in Chapter 4.
 8. No. The fact that the elasticity of demand with respect to advertising is relatively low (0.003) does not necessarily mean that an additional dollar spent on advertising would not be profitable, or that the last dollar spent was not profitable.
 9. a. -3.1 .
 b. Decreases.

- c. 2.3.
- d. 0.1.
- e. The quantity demanded will increase by 10%. (Note that Q in this problem is defined as quantity demanded *per capita*.)

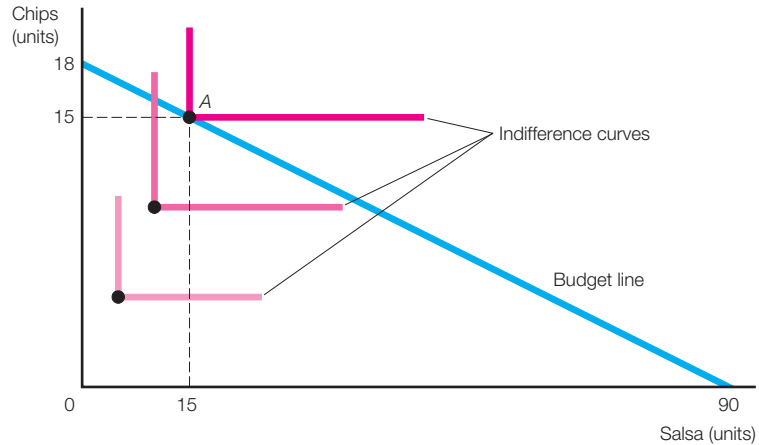
CHAPTER 3



5. His budget line is as follows:



7. Maria will maximize utility at point A, where she purchases 15 units of both chips and salsa. Note that her indifference curves are 90-degree angles.



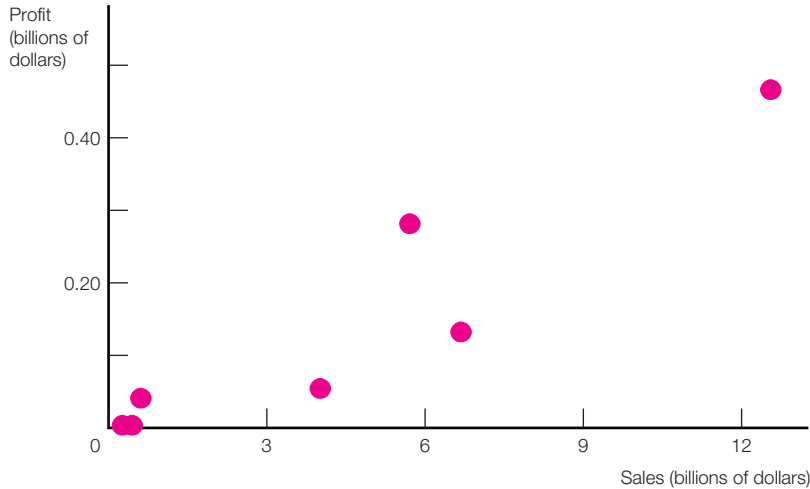
9. Since her marginal rate of substitution of opera tickets for movie tickets equals 5, and since the ratio of the price of an opera ticket to the price of a movie ticket is 10, it is impossible for her to set the marginal rate of substitution equal to the price ratio. She can increase satisfaction by substituting movie tickets for opera tickets because she is willing to give up only 5 movie tickets to get an extra opera ticket, but she has to give up 10 movie tickets to get an extra opera ticket. Thus, she will spend the entire \$300 on movie tickets; she will buy 50 of them.
11. a. 150 miles b. 300 miles c. Yes. -0.5 . d. \$3 billion.

CHAPTER 4

1. a. The evidence appears to be very strong that increases in the firm's advertising expenditure do have a positive effect on the quantity demanded of the firm's product.
- b. $Q = -104 + 3.2(5,000) + 1.5(20) + 1.6(1,000) - 2.1P$
 $= 17,526 - 2.1P$
 Thus
 $P = \frac{17,526 - Q}{2.1} = 8,346 - 0.476Q$
- c. From the answer to part b
 $Q = 17,526 - 2.1P$
 Thus, if $P = 500$
 $Q = 17,526 - 2.1(500) = 16,476$
- d. Since R^2 equals 0.89, the regression equation seems to fit the data quite well. However, we have no way of knowing (from the information given

here) whether the error terms are serially correlated or a nonlinear equation fits significantly better.

3. a. Let profit equal Y and sales equal X . Plotting Y against X , we get the following:



$$\sum X = 30.0; \sum Y = 0.94; \sum X^2 = 248.72; \sum XY = 8.307; n = 7;$$

$$\sum Y^2 = 0.3030.$$

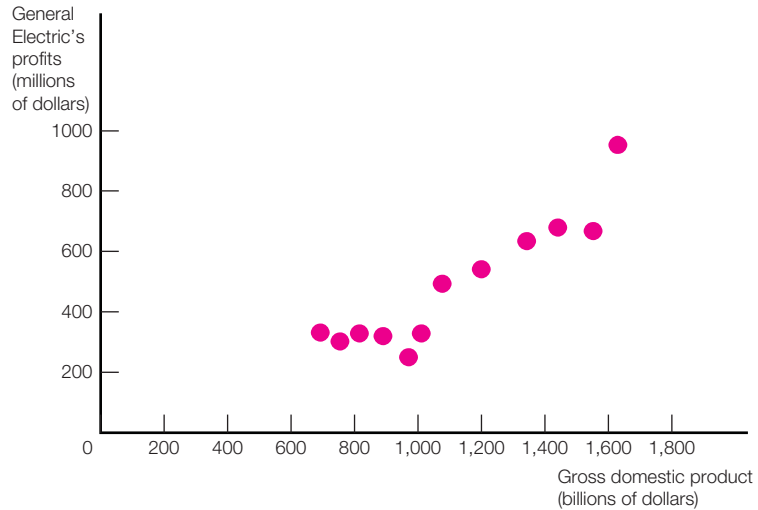
$$b = \frac{7(8.307) - (30)(0.94)}{7(248.72) - 30^2} = \frac{58.149 - 28.200}{1,741.04 - 900} = \frac{29.949}{841.04} = 0.0356$$

$$a = 0.134 - (0.0356)(4.286) = 0.134 - 0.153 = -0.019$$

The regression line is $\hat{Y} = -0.019 + 0.0356X$

- b. $-0.019 + 0.0356(2) = -0.019 + 0.071 = 0.052$. Thus, the answer is about 0.05 billion dollars.
- c. No. Prices and costs will be different in 2001 than in 1980.
5. a. 40.833.
- b. -1.025.
- c. 0.006667.
- d. 0.916.
- e. 1.361.
- f. Less than 0.001.
- g. Less than 0.001.
- h. 0.244.
- i. The average relationship is $C1 = 40.8 - 1.02 C2 + 0.00667 C3$. This relationship seems to fit the data quite well, R^2 being 0.916. There is a very small probability that the estimated effect of $C2$ (price) is due to chance, but a much higher probability (0.244) that the effect of $C3$ (disposable income) could be due to chance.

7. a. Let General Electric's profits be Y and gross domestic product be X . If we plot Y against X , we get the following graph:



To calculate a and b , we can compute the following:

	X	Y	X^2	Y^2	XY
	688	355	473,344	126,025	244,240
	753	339	567,009	114,921	255,267
	796	361	633,616	130,321	287,356
	868	357	753,424	127,449	309,876
	936	278	876,096	77,284	260,208
	982	363	964,324	131,769	356,466
	1,063	510	1,129,969	260,100	542,130
	1,171	573	1,371,241	328,329	670,983
	1,306	661	1,705,636	436,921	863,266
	1,407	705	1,979,649	497,025	991,935
	1,529	688	2,337,841	473,344	1,051,952
	1,706	931	2,910,436	866,761	1,588,286
Sum	13,205	6,121	15,702,585	3,570,249	7,421,965
Mean	1,100.42	510.08			

The results are

$$\begin{aligned}
 b &= \frac{12(7,421,965) - (13,205)(6,121)}{12(15,702,585) - 13,205^2} \\
 &= \frac{89,063,580 - 80,827,805}{188,431,020 - 174,372,025} \\
 &= \frac{8,235,775}{14,058,995} = 0.586
 \end{aligned}$$

and

$$a = 510.08 - (0.586)(1,100.42) = 510.08 - 644.85 = -134.77$$

Thus, the slope equals 0.586, and the intercept equals -134.77 millions of dollars.

- b. On the average a \$1 increase in the GDP seems to be associated with a \$0.000586 increase in General Electric's profits (recalling that GDP is measured in billions of dollars, while General Electric's profits are measured in millions of dollars).
- c. The forecast equals $-134.77 + 0.586(2,000) = -134.77 + 1,172 = 1,037.23$. That is, it equals 1,037.23 million dollars.
- d. $r^2 = 0.90$.
- e. No. No. A nonlinear relationship might be as good or better.
- f. If nothing else is available, this model may be serviceable, but it is so crude that it is difficult to believe that the analyst could not improve upon it by taking other independent variables into account.

9. a. Taking antilogs

$$Q = 102P^{-0.148}Z^{0.258}$$

$$\partial Q / \partial P = -0.148(102P^{-1.148}Z^{0.258}) = -0.148 Q / P$$

Since the price elasticity of demand equals $(\partial Q / \partial P)(P / Q)$, it follows that the price elasticity of demand equals -0.148 .

- b. $\partial Q / \partial Z = 0.258(102P^{-0.148}Z^{-0.742}) = 0.258 Q / Z$.

Since the cross elasticity of demand equals $(\partial Q / \partial Z)(Z / Q)$, it follows that the cross elasticity of demand equals 0.258.

- c. The regression seems to provide a good fit. The fact that \bar{R}^2 equals 0.98 means that 98% of the variation in $\log Q$ can be explained by the regression (see the chapter appendix). See Figure 5.6.

11. a. No.

- b. The market supply curve for wine.

CHAPTER 5

1. a. To see whether 400 hours of skilled labor and 100 hours of unskilled labor are the optimal input combination, the Elwyn Company should pick an input combination where

$$\frac{MP_S}{P_S} = \frac{MP_U}{P_U}$$

where MP_S is the marginal product of skilled labor, MP_U is the marginal product of unskilled labor, P_S is the price of skilled labor, and P_U is the price of unskilled labor. Since $P_S = 10$, $P_U = 5$, and

$$MP_S = \frac{\partial Q}{\partial S} = 300 - 0.4S$$

$$MP_U = \frac{\partial Q}{\partial U} = 200 - 0.6U$$

it follows that the Elwyn Company should pick an input combination where

$$\frac{300 - 0.4S}{10} = \frac{200 - 0.6U}{5}$$

or

$$1,500 - 2S = 2,000 - 6U$$

$$S = -250 + 3U$$

Thus, 400 hours of skilled labor and 100 hours of unskilled labor are not the optimal input combination, because, if $S = 400$ and $U = 100$, this equation does not hold.

- b. If a total of \$5,000 is spent on skilled and unskilled labor,

$$10S + 5U = 5,000,$$

since $P_S = 10$ and $P_U = 5$. From the answer to part a, we know that $S = -250 + 3U$.

Solving these two equations simultaneously, $S = 392.9$ and $U = 214.3$. Thus, to maximize output, Elwyn should hire about 393 hours of skilled labor and about 214 hours of unskilled labor.

- c. $MP_U \cdot P$ must equal P_U , where P is the price of the product. (Under present circumstances, the marginal revenue product of unskilled labor equals $MP_U \cdot P$, and the marginal expenditure on unskilled labor equals P_U .) Thus, since

$$P = 10, P_U = 5 \text{ and } MP_U = 200 - 0.6U$$

$$10(200 - 0.6U) = 5$$

$$U = 332.5$$

To maximize profit, Elwyn should hire 332.5 hours of unskilled labor. (Note that we no longer assume that a total of \$5,000 is spent on labor. Thus, the answer is different from that in part b.)

2. a. No.
- b. 50 pounds, since half of these amounts (that is, 50 pounds of hay and 125.1 pounds of grain) results in a 25-pound gain.
- c. $-(125.1 - 130.9)/(50 - 40) = 0.58$.
- d. No, because it is impossible to tell (from the information given in the question) how much hay and grain can be used to produce a 25-pound gain after the advance in technology.
4. a. No.
- b. General farms.
- c. No.
6. a&b. The average and marginal products of grain when each amount is used are calculated as follows:

Amount of grain	Average product	Marginal product
1,200	$5,917/1,200 = 4.93$	$\frac{7,250 - 5,917}{1,800 - 1,200} = 2.22$
1,800	$7,250/1,800 = 4.03$	
2,400	$8,379/2,400 = 3.49$	$\frac{8,379 - 7,250}{2,400 - 1,800} = 1.88$
3,000	$9,371/3,000 = 3.12$	
		$\frac{9,371 - 8,379}{3,000 - 2,400} = 1.65$

- c. Yes. The marginal product of grain decreases as more of it is used.
7. a. To minimize cost, the manager should choose an input combination where $MP_L/P_L = MP_K/P_K$, where MP_L is the marginal product of labor, MP_K is the marginal product of capital, P_L is the price of labor, and P_K is the price of capital. Since

$$MP_L = \frac{\partial Q}{\partial L} = 5K \quad \text{and} \quad MP_K = \frac{\partial Q}{\partial K} = 5L$$

it follows that

$$\frac{5K}{1} = \frac{5L}{2}$$

or $K = L/2$. Since $Q = 20$, $K = 4/L$. Thus

$$\frac{L}{2} = \frac{4}{L} \quad \text{or} \quad L^2 = 8$$

which means that the firm should use $2(2)^{0.5}$ units of labor and $(2)^{0.5}$ units of capital.

- b. If the price of labor is \$2 per unit, the optimal value of K is 2, and the optimal value of L is 2. Thus, output per unit of labor is $20/2$, or 10, whereas it formerly was $20/2 \times 2^{0.5}$ or $10/2^{0.5}$. Thus, output per unit of labor will rise.
- c. No, because a 1% increase in both K and L results in more than a 1% increase in Q .

CHAPTER 6

1. a. It is the cheapest of these three ways of making steel. Using this method, cost per ton is \$310.34, as compared with \$368.86 and \$401.73 with the other methods.
- b. If the price of scrap rises, the cost of producing steel based on the electric-furnace continuous-casting route will increase, because this route

uses scrap. Thus, the cost advantage of this route will be reduced if the price of scrap goes up.

- c. It suggests that U.S. steel producers may have a hard time competing with steel producers in low-wage countries.
 - d. If each figure is the minimum value of long-run average cost for a particular technique, it also equals the long-run marginal cost for the technique, since marginal cost equals average cost when the latter is a minimum.
3. a. If Q is the sales volume,
 $Q(\$200) - \$5,000 = \$10,000$,
 so Q must equal 75.
 b. Since $Q(\$250) - \$5,000 = \$10,000$, Q must equal 60.
 c. Since $Q(\$265) - \$5,000 = \$10,000$, Q must equal 56.6.
 5. The table is as follows:

Total fixed cost	Total variable cost	Average fixed cost	Average variable cost
50	0	—	—
50	25	50	25
50	50	25	25
50	70	$16\frac{2}{3}$	$23\frac{1}{3}$
50	85	$12\frac{1}{2}$	$21\frac{1}{4}$
50	100	10	20
50	140	$8\frac{1}{3}$	$23\frac{1}{3}$
50	210	$7\frac{1}{7}$	30

7. a. Yes. Since $(\partial TC / \partial Q)(Q / TC) = \alpha_1$, this is true.
 b. Yes. If $\alpha_1 < 1$, a 1% increase in output results in a less than 1% increase in total cost, so average cost falls with increases in output; in other words, there are economies of scale. If $\alpha_1 > 1$, a 1% increase in output results in a more than 1% increase in total cost, so average cost increases with increases in output; in other words, there are diseconomies of scale.

$$c. \frac{TC}{P_K} = \alpha_0 Q^{\alpha_1} \left(\frac{P_L}{P_K} \right)^{\alpha_2}$$

and

$$\log \left(\frac{TC}{P_K} \right) = \log \alpha_0 + \alpha_1 \log Q + \alpha_2 \log \left(\frac{P_L}{P_K} \right)$$

If this is treated as a regression equation, one can estimate the value, using the regression technique discussed in Chapter 4, subject to the caveats concerning various kinds of possible errors cited there.

9. a. Since marginal cost equals $dTVC/dQ$, it equals
 $MC = 50 - 20Q + 3Q^2$

It is a minimum when

$$\frac{dMC}{dQ} = -20 + 6Q = 0, \quad \text{or} \quad Q = 20/6$$

- b. Average variable cost equals

$$AVC = \frac{TVC}{Q} = 50 - 10Q + Q^2$$

It is a minimum when

$$\frac{dAVC}{dQ} = -10 + 2Q = 0, \quad \text{or} \quad Q = 5$$

- c. If $Q = 5$, average variable cost equals $50 - 10(5) + 5^2 = 25$. Marginal cost equals $50 - 20(5) + 3(5^2) = 25$. Thus, marginal cost equals average variable cost at this output level.
11. a. Using equation (6.6), $S = (23,000 + 11,000 - 30,000)/30,000 = 0.13$.
- b. Production facilities used to make one product sometimes can be used to make another product, and by-products resulting from the production of one product may be useful in making other products.

CHAPTER 7

1. a. Since average cost (AC) must be a minimum, and since

$$AC = \frac{25,000}{Q} + 150 + 3Q$$

$$\frac{dAC}{dQ} = \frac{-25,000}{Q^2} + 3 = 0$$

$$\text{Thus, } Q = \left(\frac{25,000}{3} \right)^{0.5} = 91.3, \text{ and}$$

$$AC = 25,000/91.3 + 150 + 3(91.3) = 697.7$$

so the price must be \$697.7, since in long-run equilibrium, price equals the minimum value of average cost.

- b. 91.3 units.
3. a. Marginal cost equals

$$MC = \frac{dTC}{dQ} = 4 + 4Q$$

Setting marginal cost equal to price, we have

$$4 + 4Q = 24$$

$$4Q = 20$$

$$Q = 5$$

Thus, the optimal output rate is 5.

- b. Profit equals total revenue minus total cost. Since total revenue equals $24Q$, profit equals
- $$\pi = 24Q - 200 - 4Q - 2Q^2 = -200 + 20Q - 2Q^2$$
- Because $Q = 5$
- $$\pi = -200 + 20(5) - 2(5)^2 = -200 + 100 - 50 = -150$$
- Thus, the firm loses \$150 (which is less than if it shuts down).
5. a. The White Company's marginal cost is $dTC/dQ = MC = 20 + 10Q$. Equating this to the market price $= P = 50$ and solving yields the optimal output Q , i.e., $P = 50 = 20 + 10Q = MC$ or $10Q = 30$ or $Q = 3$.
- b. The White Company's total revenue (TR) is $TR = P \cdot Q = 50 \cdot 3 = 150$. The White Company's total cost (TC) is $TC = 1,000 + 20 \cdot 3 + 5 \cdot 3 \cdot 3 = 1,000 + 60 + 45 = 1,105$. The White Company's economic profit is $TR - TC = 150 - 1,105 = -955$.
- c. The White Company's average total cost (ATC) is $ATC = TC/Q = (1,000/Q) + 20 + 5Q = (1,000/3) + 20 + 5 \cdot 3 = 333.33 + 20 + 15 = 368.33$.
- d. The industry is not in equilibrium because the firms in the industry are losing money. In the long run, we would expect some firms to leave the industry such that in the long run, the typical firm would have long run average cost equal to long run marginal cost and no economic profits being made by any firm.

CHAPTER 8

1. a. Marginal revenue $= 100 - 2Q$; marginal cost $= 60 + 2Q$. Thus, if marginal revenue equals marginal cost, $100 - 2Q = 60 + 2Q$, so $Q = 10$.
- b. Since $P = 100 - Q$, P must equal 90 if $Q = 10$. Thus, he should charge a price of \$90.
3. a. Since $P = (8,300 - Q)/2.1 = 3,952 - 0.476Q$
- $$MR = 3,952 - 0.952Q$$
- b. $MC = 480 + 40Q$. If $MC = MR$
- $$480 + 40Q = 3,952 - 0.952Q$$
- $$40.952Q = 3,472$$
- $$Q = 84.8$$
- Thus, the firm would produce 84.8 lasers per month. If $Q = 84.8$, $P = 3,952 + 0.476(84.8) = 3,912$. Thus, the price should be \$3,912.
- c. The firm's monthly profit equals
- $$84.8(3,912) - [2,200 + 480(84.8) + 20(84.8)^2] = \$145,012.80$$
5. a. If the firm is producing 5 units in the first plant, the marginal cost in the first plant equals $20 + 2(5)$, or 30. Thus, if the manager is minimizing costs, marginal cost in the second plant must also equal 30; this means that

$$10 + 5Q_2 = 30$$

$$Q_2 = 4$$

Thus, the second plant must be producing 4 units of output.

- b. Since $MC_1 = MC_2 = MC$ and the firm's output, Q , equals $Q_1 + Q_2$

$$Q_1 = (MC_1/2) - 10$$

$$Q_2 = (MC_2/5) - 2$$

$$Q = Q_1 + Q_2 = 0.7MC - 12$$

$$MC = (1/0.7)(Q + 12)$$

- c. No, because we do not have information concerning the fixed costs of each plant. But you can determine average variable cost.
7. a. It probably tended to increase because high profits induced entry. Also, the recent recession may have resulted in more demand for the services of pawnshops.
- b. No. It is likely to be an oligopoly, since there generally is not a very large number of pawnshops in a small city.
- c. Apparently not, but licensing requirements may exist.
9. a. The total revenue (TR) for diamonds is $TR_Z = P_Z \cdot Q_Z = (980 - 2Q_Z) \cdot Q_Z = 980Q_Z - 2Q_Z^2$. The marginal revenue for diamonds is $dTR_Z/dQ_Z = MR_Z = 980 - 4Q_Z$. The marginal cost for diamonds is $MC_Z = dTC/dQ_Z = 50 + Q_Z$. To maximize profit, the monopolist sets $MR_Z = MC_Z$ or $MR_Z = 980 - 4Q_Z = 50 + Q_Z = MC_Z$ or $5Q_Z = 930$ or $Q_Z = 186$. Substituting $Q_Z = 186$ into the demand function yields $P_Z = 980 - 2 \cdot 186 = 980 - 372 = 608$. Consumer surplus (CS) is then $CS_Z = 0.5 \cdot (980 - 608) \cdot 186 = 0.5 \cdot 372 \cdot 186 = 34,596$. Total revenue is $TR_Z = 608 \cdot 186 = 113,088$. Variable cost is $VC_Z = 50 \cdot 186 + 0.5 \cdot 186 \cdot 186 = 9,300 + 17,298 = 26,598$. So variable cost profit = producer surplus = $PS_Z = TR_Z - VC_Z = 113,088 - 26,598 = 86,490$. Social welfare is $CS_Z + PS_Z = 34,596 + 86,490 = 121,086$.
- b. If De Beers acts as a perfect competitor, they would set price = $P_Z = MC_Z$ or $P_Z = 980 - 2Q_Z = 50 + Q_Z = MC_Z$ or $3Q_Z = 930$ or $Q_Z = 310$. Substituting $Q_Z = 310$ into the demand function gives $P_Z = 980 - 2 \cdot 310 = 980 - 620 = 360$. Consumer surplus (CS) is then $CS_Z = 0.5 \cdot (980 - 360) \cdot 310 = 0.5 \cdot 620 \cdot 310 = 96,100$. Total revenue is $TR_Z = 360 \cdot 310 = 111,600$. Variable cost is $VC_Z = 50 \cdot 310 + 0.5 \cdot 310 \cdot 310 = 15,500 + 48,050 = 63,550$. So variable cost profit = producer surplus = $PS_Z = TR_Z - VC_Z = 111,600 - 63,550 = 48,050$. Social welfare is $CS_Z + PS_Z = 96,100 + 48,050 = 144,150$.
- c. Social welfare increases by $144,150 - 121,086 = 23,064$.
11. a. To earn 20% on a total investment of \$250,000, profit must equal \$50,000 per year. Thus, if the plant operates at 80% of capacity (and managers sell 10,000 units), managers must set a price of \$15 per unit. (Since average cost equals \$10, profit per unit will be \$5, so total profit per year will be \$50,000.)

- b. From the information given, there is no assurance that managers can sell 10,000 units per year if they charge a price of \$15 per unit.
- c. Unless the markup bears the proper relationship to the price elasticity of demand, the manager probably is sacrificing profit.
13. a. Backus' total revenues equal

$$TR = P_X Q_X + P_Y Q_Y = (400 - Q_X)Q_X + (300 - 3Q_X)Q_Y$$
and since $Q_Y = 2Q_X$

$$TR = (400 - Q_X)Q_X + (300 - 6Q_X)(2Q_X)$$

$$= 400Q_X - Q_X^2 + 600Q_X - 12Q_X^2 = 1,000Q_X - 13Q_X^2$$
Thus, the firm's profit equals

$$\pi = 1,000Q_X - 13Q_X^2 - 500 - 3Q_X - 9Q_X^2$$

$$= -500 + 997Q_X - 22Q_X^2$$
Setting $d\pi/dQ_X = 997 - 44Q_X = 0$, we find that the profit-maximizing value of $Q_X = 997/44 = 22.66$. Thus, Backus should produce and sell 22.66 units of product X and 45.32 units of product Y per period of time.
- b. The price of product X must be $400 - 22.66 = \$377.34$, and the price of product Y must be $300 - 3(45.32) = \$164.05$.

We have assumed that Backus sells all that it produces of both products. The marginal revenue of product X equals $400 - 2(22.66) = 354.68$, and the marginal revenue of product Y equals $300 - 6(45.32) = 28.09$. Since both are nonnegative, this assumption is true if Backus maximizes profit.

CHAPTER 9

1. a. The recommendation is not correct. Profit maximization requires the marginal revenue (MR) in each market be the same and equal to marginal cost. Using the relationship that $MR = P(1 + [1/\eta])$, $MR_J = P_J(1 + [1/\eta_J]) = P_J(1 + [1/-4]) = P_J(1 - [1/4]) = 0.75P_J$, $MR_{US} = P_{US}(1 + [1/\eta_{US}]) = P_{US}(1 + [1/-2]) = P_{US}(1 - [1/2]) = 0.5P_{US}$, and $MR_E = P_E(1 + [1/\eta_E]) = P_E(1 + [1/(2 - 4/3)]) = P_E(1 - [3/4]) = 0.25P_E$, where $J = \text{Japan}$, $US = \text{United States}$, and $E = \text{Europe}$. Thus, profit maximization requires $MR_J = MR_{US} = MR_E$ or $0.75P_J = 0.5P_{US} = 0.25P_E$. $0.75P_J = 0.75 \times \$1,000 = \750 , $0.5P_{US} = 0.5 \times \$2,000 = \$1,000$, and $0.25P_E = 0.25 \times \$3,000 = \750 . Since $MR_J = MR_{US} = MR_E$ does not hold, this is not a profit maximizing pricing policy.
- b. Since the US price is too high (see a. above), we should not be surprised that the sales (Q) in the US are below expectations.
- c. The decision to lower the price in the US to \$1,500 results in $MR_J = MR_{US} = MR_E$ since $MR_{US} = 0.5P_{US} = 0.5 \times 1,500 = 750$. We cannot tell

if this is a wise decision because we don't know if the marginal cost of the Ridgeway Corporation is 750.

- d. We do not know if the Ridgeway Corporation is maximizing profit because we don't know their marginal cost. Profit maximization requires $MR_J = MR_{US} = MR_E = MC$.
3. a. The firm's profit equals $P_C Q_C + P_M Q_M - TC$, or

$$\pi = (495 - 5Q_C)Q_C + (750 - 10Q_M)Q_M - 410 - 8(Q_C + Q_M)$$

 Thus

$$\frac{\partial \pi}{\partial Q_C} = 495 - 10Q_C - 8 = 0$$

$$\frac{\partial \pi}{\partial Q_M} = 750 - 20Q_M - 8 = 0$$

Consequently, $Q_C = 48.7$ and $Q_M = 37.1$, so

$$P_C = 495 - 5(48.7) = 251.5$$

b. $P_M = 750 - 10(37.1) = 379$.

- c. Yes. Under these circumstances,

$$Q_C = \frac{495 - P}{5} \quad \text{and} \quad Q_M = \frac{750 - P}{10}$$

so

$$Q = Q_C + Q_M = 174 - 0.3P$$

and

$$P = (174 - Q)/0.3 = 580 - {}^{10/3}Q$$

Thus

$$\begin{aligned} \pi &= (580 - {}^{10/3}Q)Q - 410 - 8Q \\ &= -410 + 572Q - {}^{10/3}Q^2 \end{aligned}$$

If π is a maximum

$$\frac{\partial \pi}{\partial Q} = 572 - {}^{20/3}Q = 0$$

$$\text{so } Q = 572(3/20) = 85.8$$

Consequently

$$\pi = -410 + 572(85.8) - {}^{10/3}(85.8^2) = 24,128.8$$

which compares with

$$\begin{aligned} \pi &= [495 - 5(48.7)]48.7 + [750 - 10(37.1)]37.1 - 410 \\ &\quad - 8(48.7 + 37.1) \\ &= 251.5(48.7) + 379(37.1) - 1,096.4 \\ &= 12,248.05 + 14,060.9 - 1,096.4 = 25,212.55 \end{aligned}$$

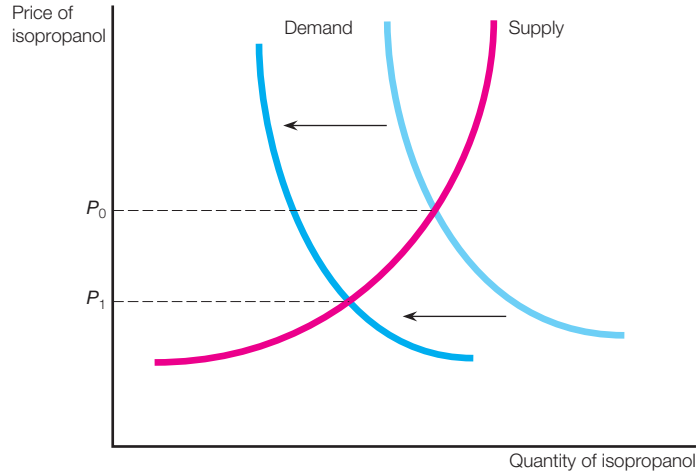
which is the value of profits when price discrimination is allowed.

So profits decrease by \$1,083.75.

CHAPTER 10

3. a. Yes. As stressed earlier, to maximize the firm's overall profit, the transfer price should equal the price of the product in the external (competitive) market.

- b. When the production of phenol increased, the supply of acetone increased, since acetone is a by-product. Thus, since less isopropanol was demanded to make acetone, the demand curve for isopropanol shifted to the left (as shown below), and the price of isopropanol declined (from P_0 to P_1).



- c. Yes.
d. Yes.

CHAPTER 11

1. a. They would want to set marginal revenue equal to the marginal cost of each firm, but this is impossible since Bergen's marginal cost is \$410 and Gutenberg's marginal cost is \$460. Because Bergen's marginal cost is always less than Gutenberg's, it will produce all the output. Equating its marginal cost to marginal revenue (MR)

$$MR = 580 - 6Q = 410$$
 so $Q = 170/6$. This is the output Bergen would produce.
 b. Nothing.
 c. Not unless Gutenberg receives an attractive share of the profit from Bergen's output even though it produces nothing.
3. a. \$9,000. b. 6.
5. a. To find the profit-maximizing price, the IATA should construct the marginal cost curve for the cartel as a whole. Then, as shown in Figure 11.2, it should determine the amount of traffic (which is the output of this industry) where marginal revenue equals marginal cost. The price that will elicit this level of traffic is the profit-maximizing price.
 b. If IATA wants to maximize profit, it will allocate this traffic among the airlines in such a way that the marginal cost of all airlines is equal. (How-

ever, for reasons discussed on page 408, it may not want to maximize profit.)

- c. No. This would not maximize profit.

7. a. Letting Alliance's profit be π_1

$$\pi_1 = Q_1[200,000 - 6(Q_1 + Q_2)] - 8,000 Q_1$$

Letting Bangor's profit be π_2

$$\pi_2 = Q_2[200,000 - 6(Q_1 + Q_2)] - 12,000 Q_2$$

If Alliance maximizes its profit, assuming that Bangor will hold its output constant

$$\frac{\partial \pi_1}{\partial Q_1} = 192,000 - 6Q_2 - 12Q_1 = 0$$

If Bangor maximizes its profit, assuming that Alliance will hold its output constant

$$\frac{\partial \pi_2}{\partial Q_2} = 188,000 - 6Q_1 - 12Q_2 = 0$$

Solving these equations simultaneously, $Q_1 = 196,000/18 = 10,888.89$, and

$$Q_2 = (188,000 - 196,000/3)/12 = 122,667/12 = 10,222.22$$

so

$$P = 200,000 - 6(10,888.89 + 10,222.22) = \$73,333.33$$

- b. Alliance's output is 10,888.89, and Bangor's output is 10,222.22.

- c. Alliance's profit is $10,888.89(73,333.33 - 8,000)$, or approximately \$711.41 million.

Bangor's profit is $10,222.22(73,333.33 - 12,000)$, or approximately \$626.96 million.

9. a. Obviously, Procter and Gamble must be concerned with its own costs. If it adopts a tactic that is far more costly to itself than to a potential entrant, it may cost more than it is worth. If the costs of the strategy outweigh the benefits, Procter and Gamble, on net, will lose.
- b. The point of these tactics is to raise the cost to a potential entrant, thus discouraging entry.
- c. Whether Procter and Gamble should have cut its price depends on whether the discount brands (and Kimberly-Clark, which had become a major rival) would cut their prices in response, and by how much. In fact, Procter and Gamble did reduce its price substantially (by 16% in the case of Luvs). According to the chairperson of Procter and Gamble, "We believe our profits are going to grow because we're going to get volume back."
- d. Yes. Procter and Gamble wanted to reduce what it regarded as improper imitation of its technology. On the other hand, firms that are sued often regard such suits as attempts to intimidate them.

11. a. The size of a firm is often measured by its total revenue. Perhaps a firm might feel that a higher total revenue would make the firm more visible to investors and customers. Also, its managers may be more interested in the growth of the firm than in profits. (However, they are likely to feel that profits should not fall below some minimum level.)

- b. To maximize its total revenue, it should set

$$\frac{d(PQ)}{dQ} = \frac{d(28Q - 0.14Q^2)}{dQ} = 28 - 0.28Q = 0$$

Thus, Q should equal 100, and P should equal \$14.

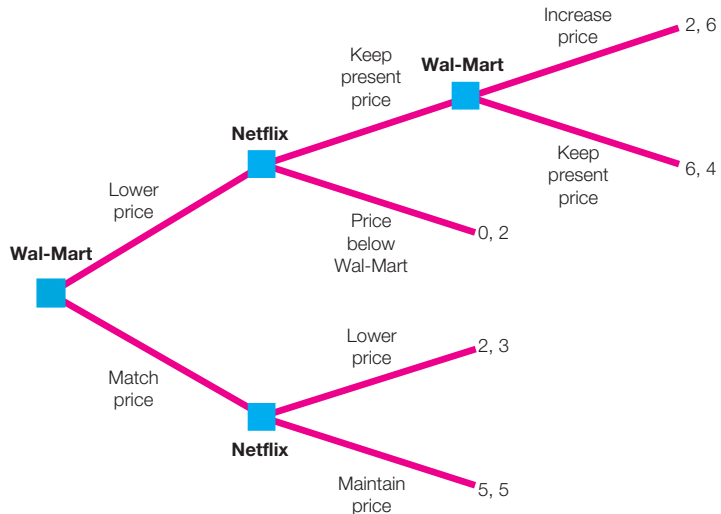
- c. If it maximizes profit, it sets

$$MR = 28 - 0.28Q = 14 = MC$$

so $Q = 50$. Consequently, the firm produces 50,000 units more than it would if it maximized profit.

CHAPTER 12

1. a. Yes. Fortnum should focus on magazines, and Maison should focus on newspapers.
b. Fortnum's profit is \$9 million, and Maison's profit is \$8 million.
c. No.
3. a.



Using backward induction, the equilibrium is for Wal-Mart to lower price, Netflix to keep its present price, and Wal-Mart to respond by keeping its present price. The final payoff to Wal-Mart is 6, and Netflix receives 4.

b.

Netflix

		Maintain Price; Price Below Wal-Mart	Maintain Price; Keep Present Price	Lower Price; Price Below Wal-Mart	Lower Price; Keep Present Price
Wal-Mart	Match Price	(W)5, 5(N)	5, 5(N)	(W)2, 3	2, 3
	Keep Present Price If Lower Price	0, 2	(W)6, 4(N)	0, 2	(W)6, 4(N)
	Increase Price If Lower Price	0, 2	2, 6(N)	0, 2	2, 6(N)

In matrix form, we find the Nash equilibria of the game's subgames. Using backward induction in (a), we found the subgame perfect Nash equilibrium.

5. a. Each firm will choose to cheat on the agreement. They will each earn \$28 million.
- b. No, as long as the horizon is finite, behavior will not change.
- c. Yes, this is an example of prisoner's dilemma.
7. If Rose cannot ascertain the strategy of its rival then it cannot implement a tit-for-tat strategy because it will not know which strategy to play. A tit-for-tat strategy requires a player to mimic the strategy played by a rival in the previous period. For example, if Rose's rival played cheat in period n , then Rose would choose cheat in period $n + 1$.

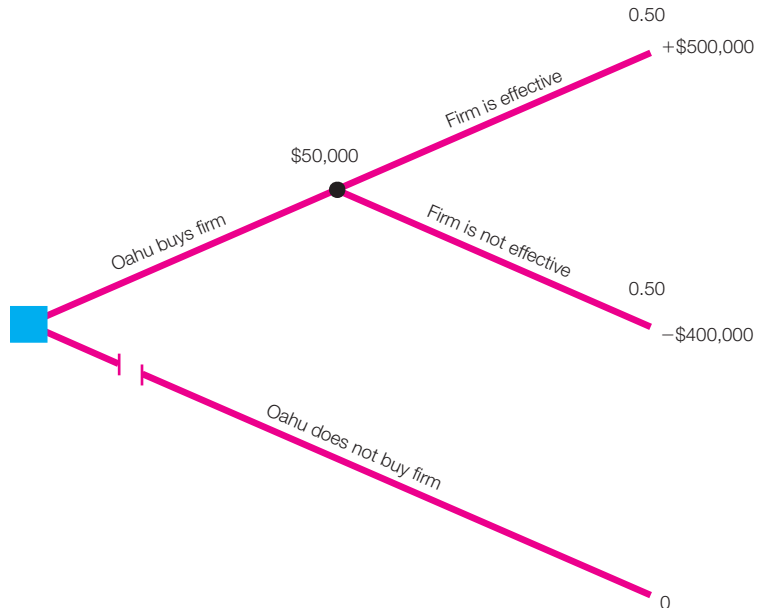
CHAPTER 13

1. Expected value or $EV(\$45) = 1,750(\$40.5)(.35) + 1,975(\$40.5)(.20) + 2,220(\$40.5)(.30) + 2,445(\$40.5)(.15)$
 $EV(\$45) = \$24,806.25 + \$15,997.50 + \$26,973 + \$14,853.38$ or $\$82,630.13$.
 $EV(\$50) = 1,200(\$45)(.35) + 1,415(\$45)(.20) + 2,001(\$45)(.30) + 2,305(\$45)(.15)$
 $EV(\$50) = \$18,900 + \$12,735 + \$27,013.50 + \$15,558.75$ or $\$74,207.25$
 $EV(\text{auction}) = \$24,806.25 + \$15,997.50 + \$27,013.50 + \$15,558.75 = \$83,376$
 So, the value of information = $\$83,376 - \$82,630.13 = \$745.87$
3. a. The mean value is 51.
- b. $z = (80 - 51)/19.85 = 1.46099$; using a standard z table, a z -score of 1.46099 is equivalent to an area under a normal distribution and to left of 80 equal to .9280. Hence the probability of a reservation price being less than 80 is approximately 93%.

5. If they choose to price their PSLs at \$6,000, their expected revenue is: $\$58,802,000 + \$58,800,000 + \$60,000,000 = \$177,602,000$. If they choose to price their PSLs at \$7,000, their expected revenue is: $\$60,214,000 + \$58,803,500 + \$56,000,000 = \$175,017,500$. If they choose to price their PSLs at \$8,000, their expected revenue is: $\$56,030,000 + \$61,610,500 + \$50,000,000 = \$167,640,500$. So, if they set the price and don't use an auction, they should charge \$6,000/PSL. If the Eagles use a modified Dutch auction, then their expected revenue is: $\$60,214,000 + \$61,610,500 + \$60,000 = \$181,824,500$. So relative to setting a price of \$6,000/PSL, the auction would increase expected revenue by \$4,222,500. But since auction costs are \$5,100,000, the Eagles are better off pricing the PSLs at \$6,000.

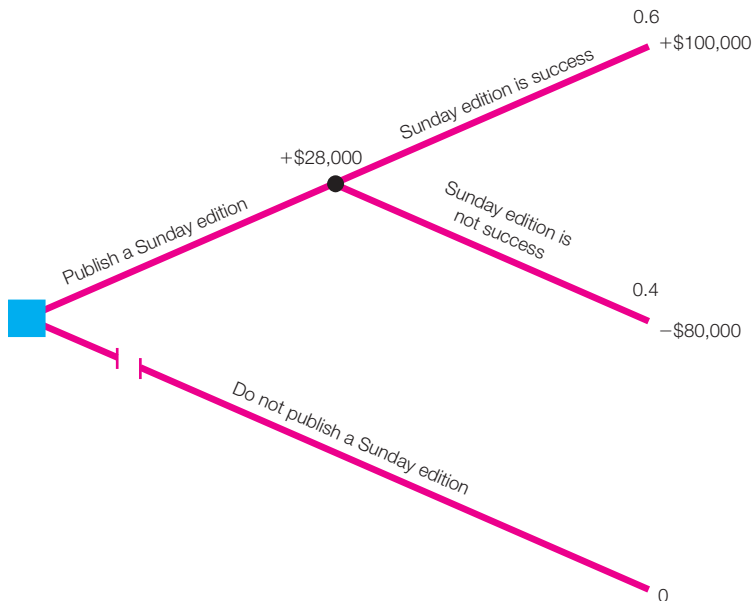
CHAPTER 14

1. a. The expected present value is \$10.7 million, the standard deviation is approximately \$5.06 million, and the coefficient of variation is 47.3%.
 - b. The expected present value is \$10 million, the standard deviation is approximately \$1.67 million, and the coefficient of variation is 16.7%.
 - c. Investment X.
 - d. Investment Y, since she is a risk averter (as indicated by the fact that U increases at a decreasing rate as P rises). Investment Y may have a lower expected present value, but it has a lower standard deviation than investment X and, more importantly, investment X's expected utility is less than investment Y's, i.e., $E(U_X) = 29.902 < 30 = E(U_Y)$.
3. No, because no probability distribution of the outcome has been given.
5. a. 3. b. -0.6 . c. -1.2 .
7. a.



- b. There is only one: whether to buy the firm or not.
- c. There is only one: whether the firm becomes an effective producer of washing machine parts or not.
- d. Yes, it should buy the firm.
- e. (1) Yes.
 (2) Three mutually exclusive outcomes are: (a) The firm becomes an effective producer of washing machine parts; (b) The firm does not become an effective producer of washing parts and is sold to the Saudis; (c) The firm does not become an effective producer of washing machine parts and cannot be sold to the Saudis.
 (3) The probability of the first outcome (in part 2) is 0.5, the probability of the second outcome is $0.5(0.2)$, or 0.1, and the probability of the third outcome is $(0.5)(0.8)$, or 0.4.
 (4) The extra profit to Oahu from the first outcome is \$500,000; the extra profit from the second outcome is \$100,000; the extra profit from the third outcome is $-\$400,000$.
- f. Oahu should buy the firm. The expected extra profit if it does so is $0.5(\$500,000) + 0.1(\$100,000) + 0.4(-\$400,000) = \$100,000$.
- g. (1) If the extra profit if the firm is made into an effective producer of washing machine parts is \$400,000 or less, the decision will be reversed. Put differently, if the *error* was an *overstatement* of this extra profit by \$100,000 or more, the decision will be reversed.
 (2) If the extra profit if the firm is made into an effective producer of washing machine parts is \$300,000 or less, the decision will be reversed. Put differently, if the *error* was an *overstatement* of this extra profit by \$200,000 or more, the decision will be reversed.

9. a.



If the publisher is risk neutral, she wants to maximize expected profit. Thus, she should publish the Sunday edition.

- b. Whether to publish the Sunday edition is a decision fork. Whether it is a success, if published, is a chance fork.

CHAPTER 15

1. a. Flat salary:

If effort is low, utility = 758.29. If effort is high, utility = 658.29. The manager chooses low effort since it gives higher expected utility. Your expected profit will be $0.3(\$5 \text{ million}) + 0.4(\$10 \text{ million}) + 0.3(\$15 \text{ million}) - \$0.575 \text{ million} = \9.425 million .

- b. Six percent of profit:

Expected utility of manager with

Low effort is $0.3\sqrt{0.06(\$5 \text{ million})} + 0.4\sqrt{1.06(\$10 \text{ million})} + 0.3\sqrt{0.06(\$15 \text{ million})} = 758.76$

High effort is $-100 + 0.3\sqrt{0.06(\$7 \text{ million})} + 0.4\sqrt{0.06(\$12 \text{ million})} + 0.3\sqrt{0.06(\$17 \text{ million})} = 736.82$

The manager chooses low effort since it gives higher expected utility. Given the manager's choice of low effort, your expected profit will be the expected profit of \$10 million minus 6% of \$10 million which is $\{0.3(\$5 \text{ million}) + 0.4(\$10 \text{ million}) + 0.3(\$15 \text{ million})\} (1 - 0.06) = \9.4 million .

- c. Five hundred thousand dollars plus half of profits in excess of \$15 million:

Expected utility of manager with

Low effort is $\sqrt{(\$0.5 \text{ million})} = 707.11$

High effort is $-100 + 0.7\sqrt{(\$0.5 \text{ million})} + 0.3\sqrt{1/2(\$17 \text{ million} - \$15 \text{ million}) + \$0.5 \text{ million}} = 762.40$

The manager chooses high effort since it gives higher expected utility. Your expected profit will be the expected profit of \$12 million minus expected compensation which is $\{\$0.5\text{m} + 0.3(0.5(\$2\text{m}))\} = \$800,000$. $\{0.3(\$7\text{m}) + 0.4(\$12\text{m}) + 0.3(\$17\text{m})\} - \$0.8\text{m} = \$11.2 \text{ million}$.

You will choose the third plan (c) since it gives highest profit after deducting the manager's compensation.

3. If the manager does not work hard, she will receive

$$[100]^{0.5} = 10$$

If the manager works hard, she will receive

$$[100 + x(1,500 - 1,300)]^{0.5} - 1$$

To calculate the minimum level of x necessary to ensure that the expected compensation will be higher with hard work, set the expected utility with hard work equal to that without hard work:

$$\begin{aligned}
[100 + x(1,500 - 1,300)]^{0.5} - 1 &= 10 \\
[100 + x(1,500 - 1,300)]^{0.5} &= 11 && \text{(square both sides to get)} \\
[100 + x(1,500 - 1,300)] &= 11^2 = 121 \\
x &= \frac{121-100}{1,500-1,300} = 0.105
\end{aligned}$$

Thus, if the manager gets a little over 10.5% of equity in excess of 1,300, she will work hard.

5. Total value of firm:

$$V = 500 + 300 - 700\left(\frac{0.2}{1+s}\right) - s = 800 - \frac{140}{1+s} - s$$

To calculate the value of s that maximizes total value, set the derivative of value with respect to s equal to zero

$$\frac{dV}{ds} = \frac{140}{(1+s)^2} - 1 = 0$$

So, $s = 10.832$.

Therefore, the value of firm is

$$800 - \frac{140}{11.832} - 10.832 = 777.34$$

Division 1 is riskless and has a stand-alone value of 500. Since all the risk comes from division 2, we must consider the chosen level of safety of this unit as a stand-alone entity. In calculating this value recollect that its value is 300 if no liability arises. However, if a loss of 700 occurs, it simply cannot pay more than the original 300 stand-alone value (because of limited liability). So, the stand-alone value of division 2 is

$$300 - 300\left(\frac{0.2}{1+s}\right) - s = 300 - \frac{60}{1+s} - s$$

Set the derivative equal to zero to maximize the division 2 stand-alone value

$$\frac{dV}{ds} = \frac{60}{(1+s)^2} - 1 = 0. \text{ So, } s = 6.746$$

So, the stand-alone value of division 2 is

$$300 - \frac{60}{7.746} - 6.746 = 285.51$$

Total value with split up

$$500 + 285.51 = 785.51$$

So the gain from split up is $785.51 - 777.34 = 8.17$

- 7.** There is an asset-substitution problem. To show this, we should value the firm as a whole, and each of the stakeholders' claims, first assuming A is chosen and then assuming B is chosen. We can then see which project selection leads to the higher value of equity. This is the one shareholders would naturally favor.

Value of the firm if project *A* is chosen:

First, note the value of the firm is either 720 (300 from existing operations and 420 from the new project) or 1020 (600 from existing operations and 420 from the new project) depending on the success of existing operations. This value must be divided up by first paying off old debt, next new debt, and finally equity.

$$\text{Value of the firm} \quad 0.5(720 + 1020) = 870$$

$$\text{Old debt} \quad 0.5(250 + 250) = 250$$

$$\text{New debt} \quad 0.5(400 + 400) = 400$$

$$\text{Equity} \quad 0.5(70 + 370) = 220$$

Value of the firm if project *B* is chosen:

The value of the firm will be either 300, 600, 1000, or 1300. These figures come from the different combinations of the two possible values for the existing operations (300 and 600) and the two values for the new project (0 and 700).

$$\text{Value of the firm} \quad 0.25(300 + 600 + 1000 + 1300) = 800$$

$$\text{Old debt} \quad 0.25(250 + 250 + 250 + 250) = 250$$

$$\text{New debt} \quad 0.25(50 + 350 + 400 + 400) = 300$$

$$\text{Equity} \quad 0.25(0 + 0 + 350 + 650) = 250$$

Shareholders would like to choose *B* after they had creditors' money. But since investors would pay only 300 for new debt, this cannot be funded. So neither project can be undertaken if debt financing is used. The analysis can be repeated using equity financing for the new project. The values of the firm will be the same as above, but these will be allocated first to the existing debt (for which 250 is owing) and any residual will accrue to equity.

Value of the firm if project *A* is chosen:

$$\text{Value of the firm} \quad 0.5(720 + 1020) = 870$$

$$\text{Old debt} \quad 0.5(250 + 250) = 250$$

$$\text{Equity} \quad 0.5(470 + 770) = 620$$

Value of the firm if project *B* is chosen:

$$\text{Value of the firm} \quad 0.25(300 + 600 + 1000 + 1300) = 800$$

$$\text{Old debt} \quad 0.25(250 + 250 + 250 + 250) = 250$$

$$\text{Equity} \quad 0.25(50 + 350 + 750 + 1050) = 550$$

Now, shareholders will naturally choose the higher net present value project *A* since it has the higher equity value. The asset-substitution problem is solved.

CHAPTER 16

1. First, note that buyers will not be willing to pay \$10,000 for any used car since there is a chance that it is a lemon. So the obvious price to contemplate is a price reflecting the average quality; i.e., there is a 75% chance the car will be "good" and worth \$10,000 and a 25% chance the car will be "bad" and worth \$5,000.

Average value $(.75)(\$10,000) + (0.25)(\$5,000) = \$8,750$

But, sellers of high quality cars, knowing their vehicles are really worth \$10,000, will not be willing to sell at this price. So, only sellers of low quality vehicles will offer their cars for sale. But buyers can anticipate that only low quality cars will be offered, therefore they will only be willing to pay \$5,000 for any secondhand Corolla on the secondhand market. Thus, only low quality cars are sold and the price is \$5,000.

3. Form the following table if all are considering buying insurance:

Group	Initial Car Value	Ending Car Value If Accident	Probability of Accident	Expected Claim
A	10,000	5,000	0.2	1,000
B	10,000	5,000	0.3	1,500
C	10,000	5,000	0.4	2,000

Thus, if all purchased insurance, the expected claims would be $1,000 + 1,500 + 2,000 = 4,500$ and hence the premium would have to be $4,500/3 = 1,500$. But this requires that ALL types buy the policy. Let's see if they will.

Would A buy insurance if the premium was 1,500? If A self insures (does not buy insurance), his/her expected utility would be

$$EU_a = 0.8(10,000)^{0.5} + 0.2(5,000)^{0.5} = 0.8(100) + 0.2(70.711) = 80 + 14.142 = 94.142$$

If A buys a full coverage policy for 1,500, A's expected utility will be

$$EU_a = (10,000 - 1,500)^{0.5} = (8,500)^{0.5} = 92.195$$

Thus, A will self insure since $94.142 > 92.195$

Thus, the premium cannot be 1,500 since A is not buying insurance.

So, if the type A's drop out of the insurance pool, we are left with the B's and C's. If only B and C are interested in buying, the expected claims are $1,500 + 2,000 = 3,500$ and the premium must be $3,500/2 = 1,750$.

But would B buy insurance at the premium of 1,750? If B self insures, their expected utility would be

$$EU_b = 0.7(10,000)^{0.5} + 0.3(5,000)^{0.5} = 0.7(100) + 0.3(70.711) = 70 + 21.213 = 91.213$$

If B buys a full coverage policy for 1,750, B's expected utility will be

$$EU_b = (10,000 - 1,750)^{0.5} = (8,250)^{0.5} = 90.830$$

Thus, B will self insure since $91.213 > 90.830$.

Thus, the premium cannot be 1,750 since B is not buying insurance.

If only C is interested in buying insurance, then the premium must be 2,000.

If C self insures, their expected utility would be

$$EU_c = 0.6(10,000)^{0.5} + 0.4(5,000)^{0.5} = 0.6(100) + 0.4(70.711) = 60 + 28.284 = 88.284$$

If C buys a full coverage policy for 2,000, C's expected utility will be

$$EU_C = (10,000 - 2,000)^{0.5} = (8,000)^{0.5} = 89.443$$

Since $89.443 > 88.284$, C will buy the insurance for a premium of 2,000.

However, No-State can sell the policy for a higher price. Type C's certainty equivalent will be $(88.284)^2 = 7,794.11$. Thus, Type C's would be willing to pay up to $10,000 - 7,794.11 = 2,205.89$ for such a policy. Since the question required that the premium must be sufficient to cover expected claims and the expected claims for type C's are

$$0.4(5,000) = 2,000, \text{ a premium of } 2,205.90 \text{ would certainly fill the bill.}$$

5. The following table charts the maximum price people are willing to pay.

Formula =

$$1.1(x)(\$50,000) \text{ for less risk averse}$$

$$1.3(x)(\$50,000) \text{ for more risk averse}$$

	Poor health Life expectancy = 9 years	Good health Life expectancy = 11 years
Less risk averse	495,000	605,000
Pay up to 1.1 times expected value		
More risk averse	585,000	715,000
Pay up to 1.3 times expected value		

At a price of \$550,000, the product will be purchased by all those in good health and the more risk averse in poor health.

The expected profit can now be calculated

$$\text{More risk averse in poor health: } (50 \text{ times } \$550,000) - (50 \text{ times } 9 \text{ times } \$50,000) = \$5,000,000$$

$$\text{Less risk averse in good health: } (50 \text{ times } \$550,000) - (50 \text{ times } 11 \text{ times } \$50,000) = \$0$$

$$\text{More risk averse in poor health: } (50 \text{ times } \$550,000) - (50 \text{ times } 11 \text{ times } \$50,000) = \$0$$

$$\text{TOTAL PROFIT} = \$5,000,000$$

CHAPTER 17

1. a. It equaled $27.5 + 21.9 + 18.5 + 9.3 = 77.2\%$. Yes.
- b. It was $27.5 + 21.9 + (18.5 + 7.3) + 9.3 = 84.5\%$ if we simply combine United's and Pan Am's shares to approximate United's postpurchase share.
3. a. If $P = 480$, $Q = 260$, according to the demand curve. Thus, the firm's total revenue equals $260(480)$ thousand dollars, or \$124,800,000. The

firm's total cost equals $50 + 0.25(260) = \$115,000,000$. Thus, the firm's accounting profit is $\$9,800,000$; this means that its rate of return is 9.8%.

- b. If it were deregulated, it would maximize

$$\begin{aligned}\pi &= (1/1,000)[Q(1,000 - 2Q)] - 50 - 0.25Q \\ &= -50 + 0.75Q - 0.002Q^2.\end{aligned}$$

Setting $d\pi/dQ = 0.75 - 0.004Q = 0$, $Q = 187.5$. Thus, under deregulation

$$\pi = -50 + 0.75(187.5) - 0.002(187.5^2) = \$20.3125 \text{ million}$$

So, the difference is $\$20.3125 - \$9.8 = \$10.5125$ million.

5. a. $220/250 = 88\%$.
- b. Yes, because it is dominated by a few firms.
- c. $225/250 = 90\%$.
- d. $140/145 = 97\%$.
7. a. No. If price is set equal to \$1, 12 firms of optimal size can exist in the market.
- b. Eight.
9. a. Because the commission tries to provide the firm with a "fair" rate of return on its investment.
- b. Because this increase reduced the firm's profit.
- c. See pages 652 to 656.

CHAPTER 18

1. a. Since the cost per patient-day (Y) = C/X , the desired relationship is

$$Y = \frac{4,700,000}{X} + 0.00013X$$

- b. To find the value of X that minimizes the value of Y , we set the derivative of Y with respect to X equal to zero

$$\frac{dY}{dX} = -\frac{4,700,000}{X^2} + 0.00013 = 0$$

Thus, $X = \left(\frac{4,700,000}{0.00013} \right)^{0.5}$ or approximately 190,141.6476 patient-days.

- c. Since $d^2Y/dX^2 = 2(4,700,000)/X^3$, d^2Y/dX^2 must be positive (since X is positive). Thus, Y must be a minimum, not a maximum, at the point where $dY/dX = 0$.
3. a. $\$5,000 - \$3,000$.
- b. 7 units per day.
- c. No, because profit is higher at 9 units per day than at 7 units per day.
5. a. Since marginal cost equals $4 + 16Q$, it is 164 when $Q = 10$.
- b. $4 + 16(12) = 196$.
- c. $4 + 16(20) = 324$.

7. a. 6
b. $24X$
c. $48X^2$
d. $8/X^3$
9. a. 2
b. $12X^2$
c. $0.8Z^{0.2}X^{-0.2} = 0.8 \frac{Y}{X}$
d. $-3Z/(4 + X)^2$
11. a. $\partial C/\partial X_1 = -3 + 4X_1 + X_2 = 0$
 $\partial C/\partial X_2 = -4 + 6X_2 + X_1 = 0$
Solving these two equations simultaneously
 $X_1 = 14/23$ and $X_2 = 13/23$
b. The answer will not change.
13. a. The Lagrangian function is $L_{TC} = 7X_1^2 + 9X_2^2 - 1.5X_1X_2 + \lambda(10 - X_1 - X_2)$. Thus
 $\partial L_{TC}/\partial X_1 = 14X_1 - 1.5X_2 - \lambda = 0$
 $\partial L_{TC}/\partial X_2 = 18X_2 - 1.5X_1 - \lambda = 0$
 $\partial L_{TC}/\partial \lambda = 10 - X_1 - X_2 = 0$
From the first two of these equations, it follows that $X_1 = (195/155)X_2$, which, together with the third equation, implies that $X_1 = 195/35$ and $X_2 = 155/35$.
b. Yes.
c. If we substitute $195/35$ for X_1 , and $155/35$ for X_2 in either of the first two equations, we find that $\lambda = 71.36$, which is the marginal cost of a rug at the cost-minimizing combination of types that total ten rugs per day.

APPENDIX A

1. a. $\frac{200,000}{10(20,000) + 0.02(50,000) + 5(10,000)} = \frac{200,000}{251,000} = 0.797$
b. $\frac{300,000}{10(30,000) + 0.02(100,000) + 5(14,000)} = \frac{300,000}{372,000} = 0.806$
c. The base year is 2000.
3. a. $\log C = 5.1 - 0.25 \log 100$
 $= 5.1 - 0.25(2)$
 $= 4.6$
Thus, $C = 39,811$.
b. $\log C = 5.1 - 0.25 \log 200$
 $= 5.1 - 0.25(2.30)$
 $= 4.525$

Thus, $C = 33,497$

- c. $1 - 33,497/39,811 = 16\%$.
5. \$9.6 million. \$7.68 million
7. a. $0.5(\$1 \text{ million}) + 0.5(\$2 \text{ million}) = \$1.5 \text{ million}$.
 b. $0.75(\$1 \text{ million}) + 0.25(\$2 \text{ million}) + \$150,000 = \1.4 million . This assumes that whether each approach costs \$1 million or \$2 million is independent of what the other approach costs. Also, the total cost figure for each approach, if adopted, includes the \$150,000. Thus only the \$150,000 spent on the aborted approach is lost. The \$150,000 spent on the approach that is adopted is part of the total cost figure given in the problem.
 c. Comparing the answers to parts a and b, parallel approaches result in lower expected cost.
9. a. $0.6(\$5 \text{ million}) + 0.4(\$3 \text{ million}) = \$4.2 \text{ million}$.
 b. $0.7(\$3 \text{ million}) + 0.3(\$5 \text{ million}) = \$3.6 \text{ million}$.
 c. $0.18(\$5 \text{ million}) + 0.82(\$3 \text{ million}) + \$500,000 = \3.86 million .

APPENDIX B

1. a. Yes. Room occupancy in August tends to be about 57% greater than in January.
 b. There are more tourists in the summer than in the winter. Because of the recession, there may have been fewer tourists, and hence the seasonal variation may have been less pronounced during the recession than before it.
 c. It might be of use in scheduling labor inputs and in ordering supplies. Certainly, the manager would want to take proper account of this seasonal variation in his or her hiring and purchasing decisions.
3. a. Because the seasonal index shows by what percent sales for a particular month tend to be above or below normal.
 b. Deseasonalized sales are as follows:

January	$2.5/0.97 = \$2.58 \text{ million}$	July	$3.2/1.01 = \$3.17 \text{ million}$
February	$2.4/0.96 = \$2.50 \text{ million}$	August	$3.1/1.03 = \$3.01 \text{ million}$
March	$2.7/0.97 = \$2.78 \text{ million}$	September	$3.2/1.03 = \$3.11 \text{ million}$
April	$2.9/0.98 = \$2.96 \text{ million}$	October	$3.1/1.03 = \$3.01 \text{ million}$
May	$3.0/0.99 = \$3.03 \text{ million}$	November	$3.0/1.02 = \$2.94 \text{ million}$
June	$3.1/1.00 = \$3.10 \text{ million}$	December	$2.9/1.01 = \$2.87 \text{ million}$

- c. Because they want to see how sales are changing, when the seasonal factor is deleted.
5. a. Yes. Yes.
 b. Yes. Yes.

7. a. $S_t = -5,744 + 2.9143t$
 b. The forecast would have been $-5,744 + 2.9143(1991) = 58.4$ billion dollars, so the forecasting error would have been about 2%.
 c. The forecast would have been $-5,744 + 2.9143(1992) = 61.3$ billion dollars, so the forecasting error would have been about 17%.
 9. a. Let $t' = 0$ when $t = 1963$. Let y be General Electric's sales.

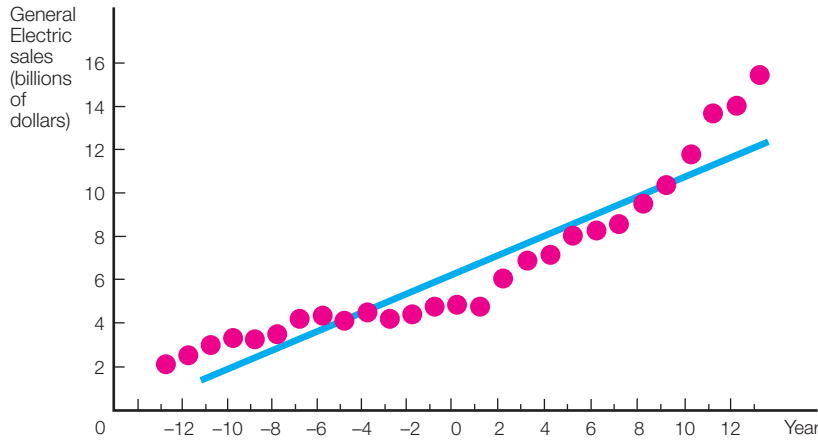
	t'	y	t'^2	y^2	$t'y$
	-13	2.2	169	4.84	-28.6
	-12	2.6	144	6.76	-31.2
	-11	3.0	121	9.00	-33.0
	-10	3.5	100	12.25	-35.0
	-9	3.3	81	10.89	-29.7
	-8	3.5	64	12.25	-28.0
	-7	4.1	49	16.81	-28.7
	-6	4.3	36	18.49	-25.8
	-5	4.2	25	17.64	-21.0
	-4	4.5	16	20.25	-18.0
	-3	4.2	9	17.64	-12.6
	-2	4.5	4	20.25	-9.0
	-1	4.8	1	23.04	-4.8
	0	4.9	0	24.01	0.0
	1	4.9	1	24.01	4.9
	2	6.2	4	38.44	12.4
	3	7.2	9	51.84	21.6
	4	7.7	16	59.29	30.8
	5	8.4	25	70.56	42.0
	6	8.4	36	70.56	50.4
	7	8.8	49	77.44	61.6
	8	9.6	64	92.16	76.8
	9	10.5	81	110.25	94.5
	10	11.9	100	141.61	119.0
	11	13.9	121	193.21	152.9
	12	14.1	144	198.81	169.2
	13	15.7	169	246.49	204.1
Sum	0	180.9	1,638	1,588.79	734.8
Mean	0	6.7			

$$b = \frac{734.8 - (180.9)(0)}{1,638 - (0)(0)} = \frac{734.8}{1,638} = 0.449$$

$$a = 6.7 - (0.449)(0) = 6.7$$

Thus, the trend is $6.7 + 0.449t'$.

b. The graph is as follows:



- c. It appears from the graph that the trend may be curvilinear, and that an exponential or quadratic trend might do better.
- d. The forecast would be $6.7 + 0.449(31) = 20.619$ billion dollars, which was only about 30% higher than actual sales in 1976. In fact, GE's sales were about \$60 billion in 1994, which indicates how poor linear extrapolations of this sort can be, particularly when one is using them to forecast many years (18 years in this case) into the future.

TABLES

TABLE E.1

<i>n</i>	Value of $\frac{1}{(1 + i)^n}$									
	Value of <i>i</i>									
	1%	2%	3%	4%	5%	6%	7%	8%	9%	10%
1	.99010	.98039	.97007	.96154	.95233	.94340	.93458	.92593	.91743	.90909
2	.98030	.96117	.94260	.92456	.90703	.89000	.87344	.85734	.84168	.82645
3	.97059	.94232	.91514	.88900	.86384	.83962	.81639	.79383	.77228	.75131
4	.96098	.92385	.88849	.85480	.82270	.79209	.76290	.73503	.70883	.68301
5	.95147	.90573	.86261	.82193	.78353	.74726	.71299	.68058	.64993	.62092
6	.94204	.88797	.83748	.79031	.74622	.70496	.66634	.63017	.59627	.56447
7	.93272	.87056	.81309	.75992	.71063	.66506	.62275	.58349	.54705	.51316
8	.92348	.85349	.78941	.73069	.67684	.62741	.58201	.54027	.50189	.46651
9	.91434	.83675	.76642	.70259	.64461	.59190	.54393	.50025	.46043	.42410
10	.90529	.82035	.74409	.67556	.61391	.55839	.50835	.46319	.42241	.38554
11	.89632	.80426	.72242	.64958	.58468	.52679	.47509	.42888	.38753	.35049
12	.88745	.78849	.70138	.62460	.55684	.49697	.44401	.39711	.35553	.31683
13	.87866	.77303	.68095	.60057	.53032	.46884	.41496	.36770	.32618	.28966
14	.86996	.75787	.66112	.57747	.50507	.44230	.38782	.34046	.29925	.26333
15	.86135	.74301	.64186	.55526	.48102	.41726	.36245	.31524	.27454	.23939
16	.85282	.72845	.62317	.53391	.45811	.39365	.33873	.29189	.25187	.21763
17	.84436	.71416	.60502	.51337	.43630	.37136	.31657	.27027	.23107	.19784
18	.83602	.70016	.58739	.49363	.41552	.35034	.29586	.25025	.21199	.17986
19	.82774	.68643	.57029	.47464	.39573	.33051	.27651	.23171	.19449	.16354
20	.81954	.67297	.55367	.45639	.37689	.31180	.25842	.21455	.17843	.14864
21	.81143	.65978	.53755	.44883	.35894	.29415	.24151	.19866	.16370	.13513
22	.80340	.64684	.52189	.42195	.34185	.27750	.22571	.18394	.15018	.12285
23	.79544	.63414	.50669	.40573	.32557	.26180	.21095	.17031	.13778	.11168
24	.78757	.62172	.49193	.39012	.31007	.24698	.19715	.15770	.12640	.10153
25	.77977	.60953	.47760	.37512	.29530	.23300	.18425	.14602	.11597	.09230

TABLE E.1 (continued)

<i>n</i>	Value of $\frac{1}{(1+i)^n}$										
	Value of <i>i</i>										
	11%	12%	13%	14%	15%	16%	17%	18%	19%	20%	24%
1	.90090	.89286	.88496	.87719	.86957	.86207	.85470	.84746	.84043	.83333	.8065
2	.81162	.79719	.78315	.76947	.75614	.74316	.73051	.71818	.70616	.69444	.6504
3	.73119	.71178	.69305	.67497	.65752	.64066	.62437	.60863	.59342	.57870	.5245
4	.65873	.63552	.61332	.59208	.57175	.55229	.53365	.51579	.49867	.48225	.4230
5	.59345	.56743	.54276	.51937	.49718	.47611	.45611	.43711	.41905	.40188	.3411
6	.53464	.50663	.48032	.45559	.43233	.41044	.38984	.37043	.35214	.33490	.2751
7	.48166	.45235	.42506	.39964	.37594	.35383	.33320	.31392	.29592	.27908	.2218
8	.43393	.40388	.37616	.35056	.32690	.30503	.28478	.26604	.24867	.23257	.1789
9	.39092	.36061	.33288	.30751	.28426	.26295	.24340	.22546	.20897	.19381	.1443
10	.35218	.32197	.29459	.26974	.24718	.22668	.20804	.19106	.17560	.16151	.1164
11	.31728	.28748	.26070	.23662	.21494	.19542	.17781	.16192	.14756	.13459	.0938
12	.28584	.25667	.23071	.20756	.18691	.16846	.15197	.13722	.12400	.11216	.0757
13	.25751	.22917	.20416	.18207	.16253	.14523	.12989	.11629	.10420	.09346	.0610
14	.23199	.20462	.18068	.15971	.14133	.12520	.11102	.09855	.08757	.07789	.0492
15	.20900	.18270	.15989	.14010	.12289	.10793	.09489	.08352	.07359	.06491	.0397
16	.18829	.16312	.14150	.12289	.10686	.09304	.08110	.07073	.06184	.05409	.0320
17	.16963	.14564	.12522	.10780	.09293	.08021	.06932	.05998	.05196	.04507	.0258
18	.15282	.13004	.11081	.09456	.08080	.06914	.05925	.05083	.04367	.03756	.0208
19	.13768	.11611	.09806	.08295	.07026	.05961	.05064	.04308	.03669	.03130	.0168
20	.12403	.10367	.08678	.07276	.06110	.05139	.04328	.03651	.03084	.02608	.0135
21	.11174	.09256	.07680	.06383	.05313	.04430	.03699	.03094	.02591	.02174	.0109
22	.10067	.08264	.06796	.05599	.04620	.03819	.03162	.02622	.02178	.01811	.0088
23	.09069	.07379	.06014	.04911	.04017	.03292	.02702	.02222	.01830	.01509	.0071
24	.08170	.06588	.05322	.04308	.03493	.02838	.02310	.01883	.01538	.01258	.0057
25	.07361	.05882	.04710	.03779	.03038	.02447	.01974	.01596	.01292	.01048	.0046

TABLE E.2

n	Value of i									
	1%	2%	3%	4%	5%	6%	7%	8%	9%	10%
1	.9901	.9804	.9709	.9615	.9524	.9434	.9346	.9259	.9174	.9091
2	1.9704	1.9416	1.9135	1.8861	1.8594	1.8334	1.8080	1.7833	1.7591	1.7355
3	2.9410	2.8839	2.8286	2.7751	2.7233	2.6730	2.6243	2.5771	2.5313	2.4868
4	3.9020	3.8077	3.7171	3.6299	3.5459	3.4651	3.3872	3.3121	3.2397	3.1699
5	4.8535	4.7134	4.5797	4.4518	4.3295	4.2123	4.1002	3.9927	3.8896	3.7908
6	5.7955	5.6014	5.4172	5.2421	5.0757	4.9173	4.7665	4.6229	4.4859	4.3553
7	6.7282	6.4720	6.2302	6.0020	5.7863	5.5824	5.3893	5.2064	5.0329	4.8684
8	7.6517	7.3254	7.0196	6.7327	6.4632	6.2093	5.9713	5.7466	5.5348	5.3349
9	8.5661	8.1622	7.7861	7.4353	7.1078	6.8017	6.5152	6.2469	5.9852	5.7590
10	9.4714	8.9825	8.7302	8.1109	7.7217	7.3601	7.0236	6.7101	6.4176	6.1446
11	10.3677	9.7868	9.2526	8.7604	8.3064	7.8868	7.4987	7.1389	6.8052	6.4951
12	11.2552	10.5753	9.9589	9.3850	8.8632	8.3838	7.9427	7.5361	7.1601	6.8137
13	12.1338	11.3483	10.6349	9.9856	9.3935	8.8527	8.3576	7.9038	7.4869	7.1034
14	13.0088	12.1062	11.2960	10.5631	9.8986	9.2950	8.7454	8.2442	7.7860	7.3667
15	13.8651	12.8492	11.9379	11.1183	10.3796	9.7122	9.1079	8.5595	8.0607	7.6061
16	14.7180	13.5777	12.5610	11.6522	10.8377	10.1059	9.4466	8.8514	8.3126	7.8237
17	15.5624	14.2918	13.1660	12.1656	11.2740	10.4772	9.7632	9.1216	8.5435	8.0215
18	16.3984	14.9920	13.7534	12.6592	11.6895	10.8276	10.0591	9.3719	8.7556	8.2014
19	17.2201	15.2684	14.3237	13.1339	12.0853	11.1581	10.3356	9.6036	8.9501	8.3649
20	18.0457	16.3514	14.8774	13.5903	12.4622	11.4699	10.5940	9.8181	9.1285	8.5136
21	18.8571	17.0111	15.4149	14.0291	12.8211	11.7640	10.8355	10.0168	9.2922	8.6487
22	19.6605	17.6581	15.9368	14.4511	13.1630	12.0416	11.0612	10.2007	9.4424	8.7715
23	20.4559	18.2921	16.4435	14.8568	13.4885	12.3033	11.2722	10.3710	9.5802	8.8832
24	21.2435	18.9139	16.9355	15.2469	13.7986	12.5503	11.4693	10.5287	9.7066	8.9847
25	22.0233	19.5234	17.4181	15.6220	14.9039	12.7833	11.6536	10.6748	9.8226	9.0770

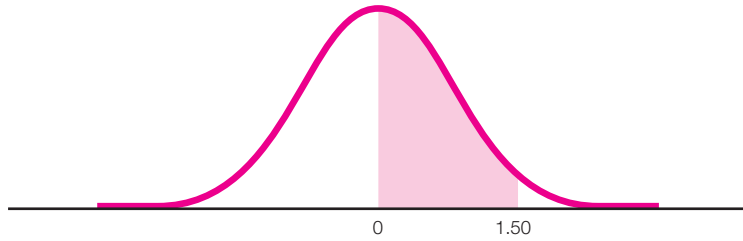
TABLE E.2 (continued)

n	Value of $\sum_{t=1}^n \frac{1}{(1+i)^t}$										
	Value of i										
	11%	12%	13%	14%	15%	16%	17%	18%	19%	20%	24%
1	.9009	.8929	.8850	.8772	.8696	.8621	.8547	.8475	.8403	.8333	.8065
2	1.7125	1.6901	1.6681	1.6467	1.6257	1.6052	1.5852	1.5656	1.5465	1.5278	1.4568
3	2.4437	2.4018	2.3612	2.3126	2.2832	2.2459	2.2096	2.1743	2.1399	2.1065	1.9813
4	3.1024	3.0373	2.9745	2.9137	2.8550	2.7982	2.7432	2.6901	2.6386	2.5887	2.4043
5	3.6959	3.6048	3.5172	3.4331	3.3522	3.2743	3.1993	3.1272	3.0576	2.9906	2.7454
6	4.2305	4.1114	3.9976	3.8887	3.7845	3.6847	3.5892	3.4976	3.4098	3.3255	3.0205
7	4.7122	4.5638	4.4226	4.2883	4.1604	4.0386	3.9224	3.8115	3.7057	3.6046	3.2423
8	5.1461	4.9676	4.7988	4.6389	4.4873	4.3436	4.2072	4.0776	3.9544	3.8372	3.4212
9	5.5370	5.3282	5.1317	4.9464	4.7716	4.6065	4.4506	4.3030	4.1633	4.0310	3.5655
10	5.8892	5.6502	5.4262	5.2161	5.0188	4.8332	4.6586	4.4941	4.3389	4.1925	3.6819
11	6.2065	5.9377	5.6869	5.4527	5.2337	5.0286	4.8364	4.6560	4.4865	4.3271	3.7757
12	6.4924	6.1944	5.9176	5.6603	5.4206	5.1971	4.9884	4.7932	4.6105	4.4392	3.8514
13	6.7499	6.4235	6.1218	5.8424	5.5831	5.3423	5.1183	4.9095	4.7147	4.5327	3.9124
14	6.9819	6.6282	6.3025	6.0021	5.7245	5.4675	5.2293	5.0081	4.8023	4.6106	3.9616
15	7.1909	6.8109	6.4624	6.1422	5.8474	5.5755	5.3242	5.0916	4.8759	4.6755	4.0013
16	7.3792	6.9740	6.6039	6.2651	5.9542	5.6685	5.4053	5.1624	4.9377	4.7296	4.0333
17	7.5488	7.1196	6.7291	6.3729	6.0472	5.7487	5.4746	5.2223	4.9897	4.7746	4.0591
18	7.7016	7.2497	6.8389	6.4674	6.1280	5.8178	5.5339	5.2732	5.0333	4.8122	4.0799
19	7.8393	7.3650	6.9380	6.5504	6.1982	5.8775	5.5845	5.3176	5.0700	4.8435	4.0967
20	7.9633	7.4694	7.0248	6.6231	6.2593	5.9288	5.6278	5.3527	5.1009	4.8696	4.1103
21	8.0751	7.5620	7.1016	6.6870	6.3125	5.9731	5.6648	5.3837	5.1268	4.8913	4.1212
22	8.1757	7.6446	7.1695	6.7429	6.3587	6.0113	5.6964	5.4099	5.1486	4.9094	4.1300
23	8.2664	7.7184	7.2297	6.7921	6.3988	6.0442	5.7234	5.4321	5.1668	4.9245	4.1371
24	8.3481	7.7843	7.2829	6.8351	6.4338	6.0726	5.7465	5.4509	5.1822	4.9371	4.1428
25	8.4217	7.8431	7.3300	6.8729	6.4641	6.0971	5.7662	5.4669	5.1951	4.9476	4.1474

TABLE E.3

Areas under the Standard Normal Curve

This table shows the area between zero (the mean of a standard normal variable) and z . For example, if $z = 1.50$, this is the shaded area shown below, which equals .4332.



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621

TABLE E.3 (continued)

Areas under the Standard Normal Curve

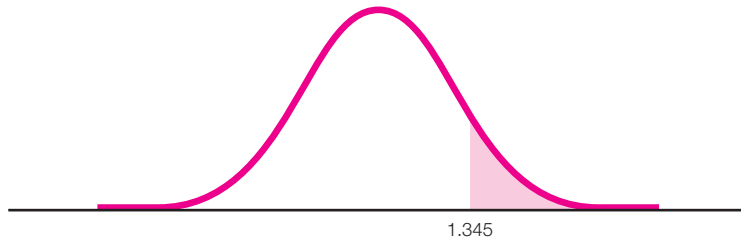
<i>z</i>	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

Source: This table is adapted from National Bureau of Standards, *Tables of Normal Probability Functions*, Applied Mathematics Series 23, U.S. Department of Commerce, 1953.

TABLE E.4

Values of t That Will Be Exceeded with Specified Probabilities

This table shows the value of t where the area under the t distribution exceeding this value of t equals the specified amount. For example, the probability that a t variable with 14 degrees of freedom will exceed 1.345 equals .10.



Degrees of freedom	Probability						
	.40	.25	.10	.05	.025	.01	.005
1	0.325	1.000	3.078	6.314	12.706	31.821	63.657
2	.289	0.816	1.886	2.920	4.303	6.965	9.925
3	.277	.765	1.638	2.353	3.182	4.541	5.841
4	.271	.741	1.533	2.132	2.776	3.747	4.604
5	0.267	0.727	1.476	2.015	2.571	3.365	4.032
6	.265	.718	1.440	1.943	2.447	3.143	3.707
7	.263	.711	1.415	1.895	2.365	2.998	3.499
8	.262	.706	1.397	1.860	2.306	2.896	3.355
9	.261	.703	1.383	1.833	2.262	2.821	3.250
10	0.260	0.700	1.372	1.812	2.228	2.764	3.169
11	.260	.697	1.363	1.796	2.201	2.718	3.106
12	.259	.695	1.356	1.782	2.179	2.681	3.055
13	.259	.694	1.350	1.771	2.160	2.650	3.012
14	.258	.692	1.345	1.761	2.145	2.624	2.977
15	0.258	0.691	1.341	1.753	2.131	2.602	2.947
16	.258	.690	1.337	1.746	2.120	2.583	2.921
17	.257	.689	1.333	1.740	2.110	2.567	2.898
18	.257	.688	1.330	1.734	2.101	2.552	2.878
19	.257	.688	1.328	1.729	2.093	2.539	2.861
20	0.257	0.687	1.325	1.725	2.086	2.528	2.845
21	.257	.686	1.323	1.721	2.080	2.518	2.831
22	.256	.686	1.321	1.717	2.074	2.508	2.819
23	.256	.685	1.319	1.714	2.069	2.500	2.807
24	.256	.685	1.318	1.711	2.064	2.492	2.797

TABLE E.4 (continued)

Values of t That Will Be Exceeded with Specified Probabilities

This table shows the value of t where the area under the t distribution exceeding this value of t equals the specified amount. For example, the probability that a t variable with 14 degrees of freedom will exceed 1.345 equals .10.

Degrees of freedom	Probability						
	.40	.25	.10	.05	.025	.01	.005
25	0.256	0.684	1.316	1.708	2.060	2.485	2.787
26	.256	.684	1.315	1.706	2.056	2.479	2.779
27	.256	.684	1.314	1.703	2.052	2.473	2.771
28	.256	.683	1.313	1.701	2.048	2.467	2.763
29	.256	.683	1.311	1.699	2.045	2.462	2.756
30	0.256	0.683	1.310	1.697	2.042	2.457	2.750
40	.255	.681	1.303	1.684	2.021	2.423	2.704
60	.254	.679	1.296	1.671	2.000	2.390	2.660
120	.254	.677	1.289	1.658	1.980	2.358	2.617
∞	.253	.674	1.282	1.645	1.960	2.326	2.576

Source: *Biometrika Tables for Statisticians* (Cambridge, U.K.: Cambridge University, 1954).

TABLE E.5

Value of an F Variable That Is Exceeded with Probability Equal to .05

	Degrees of freedom for numerator								
	1	2	3	4	5	6	7	8	9
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88

TABLE E.5 (continued)

Value of an F Variable That Is Exceeded with Probability Equal to .05

		Degrees of freedom for numerator									
		10	12	15	20	24	30	40	60	120	∞
Degrees of freedom for denominator	1	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
	2	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
	3	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
	4	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
	5	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
	6	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
	7	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
	8	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
	9	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
	10	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
	11	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
	12	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
	13	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
	14	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
	15	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
	16	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
	17	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
	18	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
	19	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
	20	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
	21	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
	22	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
	23	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
	24	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
	25	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
	26	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
	27	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
	28	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
	29	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
	30	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
	40	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
	60	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
	120	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
	∞	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

Source: *Biometrika Tables for Statisticians*.

TABLE E.6

Value of an F Variable That Is Exceeded with Probability Equal to .01

		Degrees of freedom for numerator								
		1	2	3	4	5	6	7	8	9
Degrees of freedom for denominator	1	4052	4999.5	5403	5625	5764	5859	5928	5982	6022
	2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39
	3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35
	4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66
	5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16
	6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98
	7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72
	8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91
	9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35
	10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94
	11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63
	12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39
	13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19
	14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03
	15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89
	16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78
	17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68
	18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60
	19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52
	20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46
	21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40
	22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35
	23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30
	24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26
	25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22
	26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18
	27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15
	28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12
	29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09
	30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07
	40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89
	60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72
	120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56
	∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41

TABLE E.6 (continued)

Value of an F Variable That is Exceeded with Probability Equal to .01

		Degrees of freedom for numerator									
		10	12	15	20	24	30	40	60	120	∞
Degrees of freedom for denominator	1	6056	6106	6157	6209	6235	6261	6287	6313	6339	6366
	2	99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.49	99.50
	3	27.23	27.05	26.87	26.69	26.60	26.50	26.41	26.32	26.22	26.13
	4	14.55	14.37	14.20	14.02	13.93	13.84	13.75	13.65	13.56	13.46
	5	10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
	6	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
	7	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
	8	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
	9	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
	10	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
	11	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
	12	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
	13	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
	14	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
	15	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
	16	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
	17	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
	18	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
	19	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
	20	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
	21	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
	22	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
	23	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
	24	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
	25	3.13	2.99	2.85	2.70	2.62	2.54	2.45	2.36	2.27	2.17
	26	3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23	2.13
	27	3.06	2.93	2.78	2.63	2.55	2.47	2.38	2.29	2.20	2.10
	28	3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.26	2.17	2.06
	29	3.00	2.87	2.73	2.57	2.49	2.41	2.33	2.23	2.14	2.03
	30	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80	
60	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60	
120	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38	
∞	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00	

Source: *Biometrika Tables for Statisticians*.

TABLE E.7

Values of d_L and d_U for the Durbin–Watson Test

n	A. Significance level = .05									
	$k = 1$		$k = 2$		$k = 3$		$k = 4$		$k = 5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
16	1.10	1.37	0.98	1.54	0.86	1.73	0.74	1.93	0.62	2.15
17	1.13	1.38	1.02	1.54	0.90	1.71	0.78	1.90	0.67	2.10
18	1.16	1.39	1.05	1.53	0.93	1.69	0.82	1.87	0.71	2.06
19	1.18	1.40	1.08	1.53	0.97	1.68	0.86	1.85	0.75	2.02
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
21	1.22	1.42	1.13	1.54	1.03	1.67	0.93	1.81	0.83	1.96
22	1.24	1.43	1.15	1.54	1.05	1.66	0.96	1.80	0.86	1.94
23	1.26	1.44	1.17	1.54	1.08	1.66	0.99	1.79	0.90	1.92
24	1.27	1.45	1.19	1.55	1.10	1.66	1.01	1.78	0.93	1.90
25	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	0.98	1.88
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.86
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.85
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.84
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.83
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.82
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.81
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.81
35	1.40	1.52	1.34	1.58	1.28	1.65	1.22	1.73	1.16	1.80
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.80
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.79
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.79
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.78
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
85	1.62	1.67	1.60	1.70	1.57	1.72	1.55	1.75	1.52	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
95	1.64	1.69	1.62	1.71	1.60	1.73	1.58	1.75	1.56	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78

TABLE E.7 (continued)

Values of d_L and d_U for the Durbin-Watson Test

B. Significance level = .025										
n	$k = 1$		$k = 2$		$k = 3$		$k = 4$		$k = 5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	0.95	1.23	0.83	1.40	0.71	1.61	0.59	1.84	0.48	2.09
16	0.98	1.24	0.86	1.40	0.75	1.59	0.64	1.80	0.53	2.03
17	1.01	1.25	0.90	1.40	0.79	1.58	0.68	1.77	0.57	1.98
18	1.03	1.26	0.93	1.40	0.82	1.56	0.72	1.74	0.62	1.93
19	1.06	1.28	0.96	1.41	0.86	1.55	0.76	1.72	0.66	1.90
20	1.08	1.28	0.99	1.41	0.89	1.55	0.79	1.70	0.70	1.87
21	1.10	1.30	1.01	1.41	0.92	1.54	0.83	1.69	0.73	1.84
22	1.12	1.31	1.04	1.42	0.95	1.54	0.86	1.68	0.77	1.82
23	1.14	1.32	1.06	1.42	0.97	1.54	0.89	1.67	0.80	1.80
24	1.16	1.33	1.08	1.43	1.00	1.54	0.91	1.66	0.83	1.79
25	1.18	1.34	1.10	1.43	1.02	1.54	0.94	1.65	0.86	1.77
26	1.19	1.35	1.12	1.44	1.04	1.54	0.96	1.65	0.88	1.76
27	1.21	1.36	1.13	1.44	1.06	1.54	0.99	1.64	0.91	1.75
28	1.22	1.37	1.15	1.45	1.08	1.54	1.01	1.64	0.93	1.74
29	1.24	1.38	1.17	1.45	1.10	1.54	1.03	1.63	0.96	1.73
30	1.25	1.38	1.18	1.46	1.12	1.54	1.05	1.63	0.98	1.73
31	1.26	1.39	1.20	1.47	1.13	1.55	1.07	1.63	1.00	1.72
32	1.27	1.40	1.21	1.47	1.15	1.55	1.08	1.63	1.02	1.71
33	1.28	1.41	1.22	1.48	1.16	1.55	1.10	1.63	1.04	1.71
34	1.29	1.41	1.24	1.48	1.17	1.55	1.12	1.63	1.06	1.70
35	1.30	1.42	1.25	1.48	1.19	1.55	1.13	1.63	1.07	1.70
36	1.31	1.43	1.26	1.49	1.20	1.56	1.15	1.63	1.09	1.70
37	1.32	1.43	1.27	1.49	1.21	1.56	1.16	1.62	1.10	1.70
38	1.33	1.44	1.28	1.50	1.23	1.56	1.17	1.62	1.12	1.70
39	1.34	1.44	1.29	1.50	1.24	1.56	1.19	1.63	1.13	1.69
40	1.35	1.45	1.30	1.51	1.25	1.57	1.20	1.63	1.15	1.69
45	1.39	1.48	1.34	1.53	1.30	1.58	1.25	1.63	1.21	1.69
50	1.42	1.50	1.38	1.54	1.34	1.59	1.30	1.64	1.26	1.69
55	1.45	1.52	1.41	1.56	1.37	1.60	1.33	1.64	1.30	1.69
60	1.47	1.54	1.44	1.57	1.40	1.61	1.37	1.65	1.33	1.69
65	1.49	1.55	1.46	1.59	1.43	1.62	1.40	1.66	1.36	1.69
70	1.51	1.57	1.48	1.60	1.45	1.63	1.42	1.66	1.39	1.70
75	1.53	1.58	1.50	1.61	1.47	1.64	1.45	1.67	1.42	1.70
80	1.54	1.59	1.52	1.62	1.49	1.65	1.47	1.67	1.44	1.70
85	1.56	1.60	1.53	1.63	1.51	1.65	1.49	1.68	1.46	1.71
90	1.57	1.61	1.55	1.64	1.53	1.66	1.50	1.69	1.48	1.71
95	1.58	1.62	1.56	1.65	1.54	1.67	1.52	1.69	1.50	1.71
100	1.59	1.63	1.57	1.65	1.55	1.67	1.53	1.70	1.51	1.72

TABLE E.7 (continued)

Values of d_L and d_U for the Durbin–Watson Test

n	C. Significance level = 0.01									
	$k = 1$		$k = 2$		$k = 3$		$k = 4$		$k = 5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	0.81	1.07	0.70	1.25	0.59	1.46	0.49	1.70	0.39	1.96
16	0.84	1.09	0.74	1.25	0.63	1.44	0.53	1.66	0.44	1.90
17	0.87	1.10	0.77	1.25	0.67	1.43	0.57	1.63	0.48	1.85
18	0.90	1.12	0.80	1.26	0.71	1.42	0.61	1.60	0.52	1.80
19	0.93	1.13	0.83	1.26	0.74	1.41	0.65	1.58	0.56	1.77
20	0.95	1.15	0.86	1.27	0.77	1.41	0.68	1.57	0.60	1.74
21	0.97	1.16	0.89	1.27	0.80	1.41	0.72	1.55	0.63	1.71
22	1.00	1.17	0.91	1.28	0.83	1.40	0.75	1.54	0.66	1.69
23	1.02	1.19	0.94	1.29	0.86	1.40	0.77	1.53	0.70	1.67
24	1.04	1.20	0.96	1.30	0.88	1.41	0.80	1.53	0.72	1.66
25	1.05	1.21	0.98	1.30	0.90	1.41	0.83	1.52	0.75	1.65
26	1.07	1.22	1.00	1.31	0.93	1.41	0.85	1.52	0.78	1.64
27	1.09	1.23	1.02	1.32	0.95	1.41	0.88	1.51	0.81	1.63
28	1.10	1.24	1.04	1.32	0.97	1.41	0.90	1.51	0.83	1.62
29	1.12	1.25	1.05	1.33	0.99	1.42	0.92	1.51	0.85	1.61
30	1.13	1.26	1.07	1.34	1.01	1.42	0.94	1.51	0.88	1.61
31	1.15	1.27	1.08	1.34	1.02	1.42	0.96	1.51	0.90	1.60
32	1.16	1.28	1.10	1.35	1.04	1.43	0.98	1.51	0.92	1.60
33	1.17	1.29	1.11	1.36	1.05	1.43	1.00	1.51	0.94	1.59
34	1.18	1.30	1.13	1.36	1.07	1.43	1.01	1.51	0.95	1.59
35	1.19	1.31	1.14	1.37	1.08	1.44	1.03	1.51	0.97	1.59
36	1.21	1.32	1.15	1.38	1.10	1.44	1.04	1.51	0.99	1.59
37	1.22	1.32	1.16	1.38	1.11	1.45	1.06	1.51	1.00	1.59
38	1.23	1.33	1.18	1.39	1.12	1.45	1.07	1.52	1.02	1.58
39	1.24	1.34	1.19	1.39	1.14	1.45	1.09	1.52	1.03	1.58
40	1.25	1.34	1.20	1.40	1.15	1.46	1.10	1.52	1.05	1.58
45	1.29	1.38	1.24	1.42	1.20	1.48	1.16	1.53	1.11	1.58
50	1.32	1.40	1.28	1.45	1.24	1.49	1.20	1.54	1.16	1.59
55	1.36	1.43	1.32	1.47	1.28	1.51	1.25	1.55	1.21	1.59
60	1.38	1.45	1.35	1.48	1.32	1.52	1.28	1.56	1.25	1.60
65	1.41	1.47	1.38	1.50	1.35	1.53	1.31	1.57	1.28	1.61
70	1.43	1.49	1.40	1.52	1.37	1.55	1.34	1.58	1.31	1.61
75	1.45	1.50	1.42	1.53	1.39	1.56	1.37	1.59	1.34	1.62
80	1.47	1.52	1.44	1.54	1.42	1.57	1.39	1.60	1.36	1.62
85	1.48	1.53	1.46	1.55	1.43	1.58	1.41	1.60	1.39	1.63
90	1.50	1.54	1.47	1.56	1.45	1.59	1.43	1.61	1.41	1.64
95	1.51	1.55	1.49	1.57	1.47	1.60	1.45	1.62	1.42	1.64
100	1.52	1.56	1.50	1.58	1.48	1.60	1.46	1.63	1.44	1.65

Source: J. Durbin and G. S. Watson, "Testing for Serial Correlation in Least Squares Regression," *Biometrika* 38 (June 1951).

INDEX

- actual price, 15–17
- adverse selection, 619–53, 716, 717–19
 - in annuities, 627, 629–33, 648
 - in automobile insurance, 622–26, 648
 - Excel exercise on, 650–53
 - in the job market, 639–40
 - life insurance markets and, 631–33
 - moral hazard and, 636–37, 644–47
 - problems on, 621, 642–43, 644–47, 648–50
 - in the product market, 640–43
 - self-selection and, 634–38
 - simple, 635–38
 - in used car market, 619, 620–22, 648
- advertising
 - brand loyalty and, 288–89, 290
 - price elasticity of demand and, 57–59, 288–89, 290
 - for profit maximization, 285–88, 290
- advertising agencies, economies of scope in, 192–93
- Airbus, 410, 430, 697–98
- aircraft industry, output and costs in, 183
- airlines, price discrimination and, 45–46, 304, 316–18
- airport security lines, 218–19
- Akerlof, George, 620, 634
- Albaugh, Jim, 430
- Alcoa (Aluminum Company of America), 677–78, 720
- Algeria, 420–21
- Allegheny Power, 325
- Allen, Paul G., 4
- allocation of output among plants, 298–301
- Aluminum Company of America (Alcoa), 677–578, 720
- American Airlines, 675
- American Telephone and Telegraph Company (AT&T), 720
- American Water, 226
- Amtrak, income elasticity of demand and, 54
- amusement parks, 313, 316
- Anastrozol, 684
- Anna's Linens, 512
- annuities market and adverse selection, 627, 629–33, 648
- antitrust policy, 411, 659–60, 667–74, 720
 - China and, 679
 - Clayton Act and, 676–77, 720
 - Federal Trade Commission Act and, 676, 677
 - interpretation of antitrust laws, 677–79
 - Microsoft and, 656
 - Sherman Act and, 674–75, 676, 720
- arc elasticity of demand, 39–40
- Arimidex, 684
- ARMs, 717
- art, scientific pricing for, 335
- artificial intelligence, 145
- ascending-bid (English) auctions, 504–5, 507, 508, 513–15, 523
- ascending bid timed auctions, 504
- asset substitution in principal-agent issues, 595, 599, 602–3, 606, 607
- assurance games, 491–192
- asymmetric information
 - annuities markets with, 629–31
 - in automobile insurance market, 624–26
 - in pollution cleanup, 628
- AT&T (American Telephone and Telegraph Company), 720
- Au, Miu Hing Alice, 596, 597
- auctions, 196, 501–29
 - bidding strategies in, 507–10, 524
 - collusion in, 523
 - common-value, 506
 - consumer preference-revealing nature of, 515–18
 - design of, 522–23
 - Dutch (descending-bid), 505, 507, 508, 523, 524
 - English (ascending-bid), 504–5, 507, 508, 513–15, 523
 - Excel exercise on, 526–29
 - expanding use of, 503
 - history of, 502
 - number of bidders in, 519–20
 - online, 502
 - private-value, 506, 523
 - problems on, 524–26
 - revenue generation and mechanisms of, 505–7
 - reverse, 512–13, 516–17
 - risk aversion and, 518
 - sealed-bid, 505, 507, 508
 - second-price, 505, 508, 523
 - sellers' strategies in, 510–15, 524
 - types of, 504–5
 - winner's curse and, 520–22, 524
- Australia, 468, 713
- automobiles
 - insurance for, adverse selection in, 622–26, 648
 - Model T, 779–80
 - output, forecasting, 122
 - used, 619, 620–22, 648
- average costs, short-run, 177–79, 180, 181–82
- average fixed cost, short-run, 177, 183
- average product, 140, 141, 142, 143, 181
- average profit, 730
- average total cost, short-run, 179
- average values, 731–34
- average variable cost, short-run, 177–79, 181–82
- Avon, 34
- backward induction, 479–83
- bait and switch in principal-agent issues, 606–7
- banking
 - government regulation of, 656
 - production function for, 149
- barriers to entry, 226
- baseball, 8–10, 12, 155, 411, 503
- Beautiful Mind*, A, 427
- beer, price elasticity of, 51
- Benckiser, Reckitt, 680
- Benoliel, Joel, 329

INDEX

- Berkshire Hathaway, 193
- Berman, Dennis, 469
- Bertrand model, 435–37
- Best Buy, 31, 624
- bidding
 - number of bidders and, 519–20
 - risk aversion and, 518
 - strategies for, 507–10, 524
 - winner's curse and, 520–22, 524
- Blackstone Discovery, 145
- Blair, Thomas, 547
- Blanfein, Lloyd, 605
- Blatz, 678
- blood tests, forecasting demand for, 801
- Bloomberg, Michael R., 4, 218
- Bloomberg Businessweek*, 586, 791
- Boeing Company, 6, 410, 430, 696–98
- Bono, 4
- Boston Gas Company, 666
- Bowring, Joseph, 237
- brand equity, advertising and, 288–89, 290
- Branson, Richard, 4
- Braun, Michael, 588, 589
- Brazil, 686, 696
- break-even analysis, 201–3, 204–5, 211–14
- break-even points, 201
- Breedlove Foods, 682
- Briend, Andree, 682, 683
- British Air, 219, 688
- British Office of Fair Trading, 688
- British Petroleum, 193
- Brynjolfsson, Eric, 784
- BskyB, 191
- budget line, 70–73, 85, 93
 - calculation of, 71–72
 - definition of, 71
 - equilibrium market bundle and, 73–75
 - shifts in, 72
- Buffett, Warren, 4, 262, 418, 432, 593
- bundling, 357–58, 398
 - carbon credits with gas sales, 359
 - of college textbooks, 361
 - exclusion in, 366, 367–68, 370
 - extraction in, 366
 - inclusion in, 366
 - mechanics of, 358–77
 - mixed, 357–77
 - New Yorker* and, 383
 - as a preemptive entry strategy, 381–85
 - simple, 357–58
 - tying and, 386–88
- business fluctuations, 794, 802
- business forecasting, *see* forecasting
- Calaycay, Justino B., Jr., 433
- call options, 591–92, 594, 604–5
- Canada, 683, 685, 696
- carbon credits, 359
- Cargill, 193
- Carolina Panthers, 328
- cars, *see* automobiles
- cartels, 411–14, 420–21, 423, 428, 440
- cashiers, 152
- category killers, 679
- cause marketing, 4
- CCX (Chicago Climate Exchange), 710
- CDOs (collateralized debt obligations), 716, 717–18
- Celler-Kefauver Antimerger Act, 677
- CEMCO, 808
- Census of Manufacturing*, 667
- centipede game, backward induction and, 480–81
- CEO pay, 586–87
- CERCLA (Comprehensive Environmental Response, Compensation, and Liability Act), 628
- certainty equivalence in insurance, 554–57
- certainty equivalent approach, 553–54
- CFFC (Citizens First Financial Corporation), 515–18
- chain rule, 744
- Charlotte, N.C., 328
- Chicago Board of Trade, 708
- Chicago Climate Exchange (CCX), 710
- China, 214–15, 237, 679, 683, 685, 686, 688
- Chun, David, 586
- Cicero, 502
- cigarette industry, 43
- Citigroup, 149
- Citizens First Financial Corporation (CFFC), 515–18
- Clark, Kim, 767
- Clayton Act, 676–77, 720
- Clean Air Act (1990), 708
- Coase, Richard, 708
- Cobb-Douglas production function, 164, 165
- Coca-Cola Company, 193, 593
- coefficient of determination, 109, 109n4, 110, 111, 132
 - multiple, 114, 114n7
 - variation and, 131–33, 131n14
- coefficient of variation, 553, 558
- coffee cartel, 421
- coincident series, 806
- Coke, 522
- collateralized debt obligations (CDOs), 716, 717–18
- college textbooks, bundling, 361
- Collins, Bob, 713
- collusion, 420–23, 523
- collusive agreements
 - breakdown of, 413–14
 - oligopoly, 411–13, 440
- commitment in game theory, 482–83
- common-value auctions, 506
- competition
 - in duopolies, 417–35, 435–38
 - monopolistic, *see* monopolistic competition
 - monopoly vs., 658–60
 - perfect, *see* perfect competition
- Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA), 628
- Conference Board, 807
- congestion pricing, 218
- conglomerates, 194
- constant-cost industries, 247–49, 252
- constant-elasticity demand function, 59–60, 61
- constant returns to scale, 158–60
- constants, derivatives of, 738
- constrained optimization, 752–54, 758
- consumer behavior, 65–96
 - budget line and, 70–73, 79–84, 85, 93
 - corner solutions and, 76, 78
 - equilibrium market bundle and, 73–75, 76, 85
 - indifference curves and, 66–70, 73–74, 76, 77, 85–88, 90–93
 - individual demand curve and, 84–85, 88–90
 - managers' influence on, 79–84
 - marginal rate of substitution and, 68–69, 76, 93
 - problems on, 86–87, 93–96
 - time constraints and, 75
 - utility and, 69–70, 76, 77, 79–84

- consumer goods, product liability law and, 607–12
- consumer income levels, 30–31, 32
- consumer interviews, 100–101, 123–24
- Consumer Protection Act, 657
- consumer's reservation price, 90
- consumer surplus, 90, 92, 93, 305, 691–94, 695
- consumer surveys, 100
- consumer tastes, 30, 32
- cooperative behavior, oligopolistic industries and, 411–13, 440
- coordination games, 489–93
- copper
 - actual price of, 16
 - demand curve for, 11, 13, 17–18
 - equilibrium price of, 14–15, 16–17, 19–20
 - supply curve for, 13–14, 18–20
- corner solutions, 76, 78, 156, 157
- Costco, 327, 329, 684
- cost functions, 174, 186–87, 204
 - see also* long-run cost functions; short-run cost functions
- cost-of-service model, 660–64
- cost-plus pricing, 267–73, 289
 - case studies of, 269–71
 - definition of, 267
 - elasticity and, 46, 290
 - government regulations and, 271
 - at internet companies, 269–70
 - profit maximization and, 271–73
- costs, 172–221
 - average, 177–79, 180, 181–82
 - break-even analysis and, 201–3, 204–5, 211–14
 - Excel exercise on, 209–10
 - explicit vs. implicit, 173
 - fixed, 174–77
 - historical, 173
 - incremental, 756–58
 - marginal, *see* marginal cost
 - negotiation, 195
 - network economies and, 198–201
 - opportunity, 173–74, 200, 201, 204
 - output and, 183, 230
 - of pollution and pollution control, 704–7
 - problems on, 183, 186–87, 205–9
 - profit contribution analysis and, 203–4
 - reverse auctions to reduce, 512–13
 - sunk, 174
 - total fixed, 175–77, 183
 - total variable, 175–76
 - transaction, 194–97
- coupons in price discrimination, 318–19
- Cournot analysis, 423–31, 439, 440–41, 449–57
- cranberries, market for, 414
- Crandall, Robert, 675
- credibility of the bundle, 362
- cross-price elasticities of demand, 55–57, 58, 61
- crow's feet method, 378n7
- cruise ships, 191
- cyclical variation, 802–3, 812
- D'Agostino Supermarkets, 46, 47, 335
- Dairy Farmers of America, 678
- Dauman, Philippe, 587
- deadweight loss, 662, 698, 703
- Dean Foods Co., 678
- decision forks, 557
- decision trees, 538–41, 557
- decreasing returns to scale, 158–60
- deductible, definition of, 634
- Delves, Donald, 587
- demand and supply, principles of, 7–21
 - actual price, 15–17
 - in baseball, 8–10
 - demand curve, shift in, 17–18, 21
 - demand side of market, 11–13
 - equilibrium price, 14–15, 21
 - Excel exercise on, 24–26
 - in import and export, 688–91
 - supply curve, shift in, 18–20
 - supply side of market, 13–14, 21
- demand curve, 11
 - individual, 84–85, 88–90
 - kinked, 439–40, 441
 - market, *see* market demand curve
 - shift in, 17–18, 21, 229
 - two-part tariffs and, 331, 333–39
- demand function
 - constant-elasticity, 36, 37, 59–60, 61
 - definition of, 11
 - estimating, *see* estimating demand functions
 - market, *see* market demand function
- demand theory, 28–64
 - advertising elasticity of demand and, 57–59
 - constant-elasticity and unitary elastic demand function in, 59–60, 61
 - cross-price elasticities and, 55–57, 58, 61
 - emerging markets and, 34
 - income elasticity of demand and, 53–55, 56, 61
 - market demand curve in, 29–32, 60
 - price elasticity of demand and, 28–29, 36–53, 61
 - see also* price elasticity of demand
 - problems on, 61–64
- Demby, Nicki, 596
- derivatives, 734–37, 758
 - finding, 738–44
 - to solve maximization and minimization problems, 744–48, 758
- descending-bid (Dutch) auctions, 505, 507, 508, 523, 524
- deviation from trend, definition of, 796
- DHL, 475–77
- differences, derivatives of sums and, 740–43
- diffusion models, 780–85
- diminishing marginal returns, law of, 143
- Dimon, James, 605
- direct market experiments, 101, 102
- direct selling, 34
- discounting and present values, 821–27
- discount rates, risk-adjusted, 556
- discrimination, price, *see* price discrimination
- diseconomies of scale, 191
- Disney, 193, 196, 313
- Disney Magic*, 503
- disutility of effort, 574
- Dodd–Frank Wall Street Reform, 657
- dominant strategies, 468–72, 469n1, 494–95, 507–8
- Dominicks, 47
- Douglas Aircraft, 780
- Drew, Jill, 717
- driving sales, 47
- duopoly, 410–11, 417–35
 - collusion in, 420–23
 - differentiated products and, 435–38, 441
 - quantity competition in, 423–31
 - Stackelberg behavior and, 431–35, 441, 457–59
 - Warren Buffet and, 418

INDEX

- DuPont, 774
- Dupuis, Mike, 624
- Durbin-Watson test, 125
- Dutch (descending-bid) auctions, 505, 507, 508, 523, 524
- dynamic pricing, 9–10
- Eastman Kodak, 677
- EBay, 478, 504, 619
- econometric models, 807–9, 812
- economic forecasting, *see* forecasting
- Economic Times, The* (of India), 12
- economies of scale, 189–91, 411
- economies of scope, 3, 191–94, 204
- Economist*, 47
- e-discovery software, 145
- education as a signal of labor market skills, 639–40
- efficiency
 - of regulated monopolies, 666–67
 - vs. risk sharing in executive compensation, 581–89
- effluent fees, 707–8
- effort in principal-agent issues, 572–74
 - nonobservable
 - incentive conflict and, 579–81
 - risky output and, 578–79, 581–90
 - observable, incentive conflict and, 575–79
- Eisner, Michael, 600
- Ek, Brian, 516, 517
- elasticity of demand
 - advertising, 57–59
 - arc, 39–40
 - cross-price, 55–57, 58, 61
 - definition of, 28–29, 36
 - income, 53–55, 56, 61
 - point, 39–40
 - price, *see* price elasticity of demand
 - unitary, 36, 37, 59–60
- electric power industry, 325
- Ellison, Larry, 4
- Embraer, 430
- emerging markets, servicing demands in, 34
- emissions permits, transferable, 708
- emissions trading scheme (ETS), 468, 710
- Enders, Tom, 430
- engineering data, 221
- English (ascending-bid) auctions, 504–5, 507, 508, 513–15, 523
- environmental pollution, regulation of, 703–14, 721
 - CERCLA and, 628
 - cost increases from, 710–14
 - external economies/diseconomies and, 703–4, 721
 - forms of regulation in, 707–9
 - genesis of pollution problem, 704
 - optimal level of control in, 704–7
- Equilar, 586–87
- equilibria in game theory, 467
- equilibrium
 - definition of, 467
 - long-run, 245–47, 248
 - separating, 638
 - in used car market, 621
- equilibrium market bundle, 73–75, 76, 85
- equilibrium price, 14–15, 16–17, 18–20, 21, 228–29
- Equity Seats Rights (ESR), 334
- error terms, 104
 - residuals and, 121
 - serial correlation of, 119, 120
- estimated regression coefficient, 107, 124
- estimated regression line, 105–6
- estimating demand functions, 97–133
 - coefficient of determination in, 109, 109n4, 110, 111, 131–33, 131n14
 - consumer interviews in, 100–101, 123–24
 - identification problem in, 97–100, 123–24
 - market experiments in, 101, 102
 - method of least squares in, 106–8, 106n3, 124
 - multicollinearity and, 118–19, 125
 - multiple regression and, 109, 112–13, 112n5, 113n6, 125
 - problems on, 122, 125–31
 - regression analysis in, 101–3, 113–14, 114n7, 115, 124
 - residuals in, 120–21, 123
 - sample regression line and, 105–6
 - serial correlation and, 119–20, 119n10
 - simple regression models in, 103–4
 - statistical software output and, 114–18, 115n8, 118n9
- ETS, *see* emissions trading scheme (ETS)
- European Union, 660, 688–89, 696, 720
- executive pay in Europe and Asia, 596–98
- exercise price, 594
- expansion phase of business fluctuations, 802
- expected utility, 533, 546, 562–66
- expected value, 537, 540–42, 543, 800
- Expedia, 318
- experience goods, 641
- explicit costs, 173
- exponential smoothing and forecasting, 817–19
- extensive form, 464
- external diseconomies, 704, 721
- external economies, 703–4, 721
- Exxon Mobil, 191, 193, 410
- face value, definition of, 599
- factory performance, tracking, 767–68
- Fannie Mae, 719
- Federal Communications Commission (FCC), 659, 713
- Federal Trade Commission Act, 676, 677, 720
- FedEx, 475–77, 544–45
- Financial Crisis Inquiry Commission, 657
- financial crisis of 2008, 586–87, 596, 604, 611, 657, 716–19
- Financial Times*, 18, 410
- Fingleton, John, 680, 688
- firm, theory of the, 3–5, 20
- first-degree price discrimination, 196, 306–11, 329, 339
- first-mover games, 492
- first movers, *see* market leaders
- First Solar Inc., 16
- fixed costs, 174–77
- focus groups, 100
- folk theorem, 485
- food supply, unrest over cost of, 18–19
- Ford, Henry, 779–80
- Ford Motor Company, 780
- forecasting, 54–55, 239, 791–819
 - auto output, 122
 - cyclical variation and, 802–3, 812
 - demand for blood tests, 800
 - diffusion models in, 780–85
 - econometric models in, 807–9, 812
 - elementary techniques in, 804–5

- exponential smoothing and, 817–19
- leading indicators and, 806–7, 812
- linear trends, 795–97
- nonlinear trends, 797–98
- numerical example of, 809–10
- problems on, 801, 808, 813–16
- residuals and, 810–12
- seasonal variation and, 798–802, 812
- survey techniques in, 792–93, 812
- time series in, 793–95, 812
- foreign trade, 682–83, 685–88
- Fortune Brands, 193
- franchisers and franchisees, conflicting objectives of, 266
- Freddie Mac, 719
- frequency definition of probability, 534, 536
- F* statistic, 116–17, 124–25
- full-cost pricing, *see* cost-plus pricing
- full-employment level, 802
- full insurance, definition of, 634
- functional relationships, 728–29, 758
- function of a function, derivatives of, 744
- Galenson, David, 335
- games
 - assurance, 491–192
 - coordination, 489–93
 - first-mover, 492
 - incomplete information, 485–87
 - matching, 490
 - repeated, 483–85, 495
 - strictly competitive, 493–94
 - tit for tat in, 486
 - zero-sum, 494
- game theory, 196–97, 460–500
 - backward induction in, 479–83
 - bait and switch in, 606–7
 - commitment in, 482–83
 - dominant strategies in, 468–72, 469n1, 494–95
 - equilibria in, 467
 - Excel exercise on, 498–500
 - five parameters of, 462–63, 464
 - folk theorem in, 485
 - game trees in, 464–67
 - managerial uses of, 468–69
 - Nash equilibrium in, 472–78, 488, 489
 - prisoner's dilemma in, 483–85, 488
 - problems on, 488, 495–98
 - reputation building in, 487, 489
 - solution concepts in, 467
 - strategy in, 460–63, 494
 - visual representation of, 463–67
- game trees, 464–67
- gas sales, 359
- Gates, Bill, 4, 656
- Gaviscon Advance, 680
- Gazprom, 359
- Genco Exploration, 540–41, 547, 557
- General Electric, 193–94, 428, 430
- Germany, 215, 523, 636–37
- Gilanya, 49
- Ginnie Mae, 719
- Giving Pledge, 4
- Gladstone, Alan, 512
- GlaxoSmithKline (GSK), 193, 523
- Goings, Rick, 34
- Google, 152
- government-business relations, 656–727
 - efficiency and, 666–67
 - environmental pollution and, *see* environmental pollution, regulation of
 - Excel exercise on, 725–27
 - financial crisis of 2008, 657, 716–19
 - monopoly regulation and, 658–64
 - patent system and, 680–81, 684, 720
 - price ceilings/floors and, 698–702, 720–21
 - problems on, 665, 686–87, 713, 721–25
 - public goods and, 714, 718–20, 721
 - public utilities regulation and, 664–66, 720
 - strategic trade policy and, 696–98
 - see also* antitrust policy; trade and trade policy
- Greece, ancient, 502
- greenhouse gas emission, 710–11
- Gulf Power of Pensacola, 325
- haggling, 196, 307
- Harley-Davidson, 111
- Hayes, Robert, 767
- health care debate, national
 - adverse selection and, 626
 - moral hazard and, 608
- Herbalife, 34
- Herfindahl-Hirschman index (HHI), 668–74
- Herodotus, 502
- HHI, *see* Herfindahl-Hirschman index (HHI)
- Hickey, Jim, 155
- historical cost, 173
- Hitt, Lorin, 784
- Hi Vision Ltd., 713
- Hoch, Steve, 329
- Home Depot, 414–15
- Humphrey Corporation, 273, 287
- Huntsman, Jon, 4
- IBM Corporation
 - antitrust suit against, 720
 - parallel development efforts at, 773
 - tying by, 386–88, 676
- Icahn, Carl, 4
- identification problem, 97–100, 123–24
- IIG, *see* incomplete information games (IIG)
- implicit costs, 173
- imports, U.S., 688–91
- InBev, 51
- incentive compatibility, 508, 581
- incentive conflict in principal-agent issues
 - nonobservable effort and, 578–89
 - observable effort and, 575–79
- incentive pay, 600–601
- income elasticity of demand, 53–55, 56, 61
- incomplete information games (IIG), 485–87
- increasing-cost industries, 249–51, 252
- increasing returns to scale, 157–60
- incremental costs, 756–58
- incremental revenues, 756–58
- indexed stock options, 592–94
- indifference curves, 66–70, 73–74, 76, 77, 85–88, 90–93
- individual demand curve
 - derivation of, 84–85, 88
 - market demand curve and, 88–90
- induction, backward, 479–83
- ineffective hedge, 555
- inelastic demand, 36, 37, 42
- inferior goods, 53–54

INDEX

- information technology and productivity growth, 784
- Initial Public Offerings (IPOs), 588–89
- ink, expense of, 387
- innovation, 774, 786
 - diffusion models and, 780–82
 - production function and, 149, 152
 - profit and, 6, 21
 - time-cost trade offs in, 774–77, 786
- inputs
 - fixed, 174
 - optimal combination of, 152–56, 168–71
 - substitution of, in baseball, 155
 - variable, *see* variable inputs
- insurance
 - certainty equivalence and, 554–57
 - full, 634
 - moral hazard and, 594–95
 - see also specific types of insurance*
- intercept of the regression line, 105
- internal rate of return, 826–27
- International Harvester, 677–78
- International Trade Commission (ITC), 111
- Internet
 - as an information source, 619
 - auctions on, 502
 - cost-plus pricing on, 269–70
 - development of services for, 478
 - two-part tariffs and, 327
- Interstate Commerce Commission, 659
- invisible hand, definition of, 17
- iPad2, 268
- IPOs, *see* Initial Public Offerings (IPOs)
- iron ore market, 237
- isocost curves, 153, 154, 278–81
- isoquants, 146, 148, 150–51, 162–63, 165, 764
- isorevenue lines, 278–81, 290
- iso-utility curves, *see* indifference curves
- ITC, *see* International Trade Commission (ITC)
- ITT, 193

- Japan, 394, 683, 685, 686
- Japanese auctions, 504
- JA Solar Holdings Co., 16
- Jefferies, Michael, 586
- Jerrold Electronics Corporation, 388
- job market, adverse selection in, 639–40
- joint products, pricing of
 - with fixed proportions, 274, 276–77
 - with variable proportions, 277–81, 290
- JPMorganChase, 611, 717
- Juncker, Jean-Claude, 596
- Justice Department, U.S., 422–23, 656, 675, 677, 688

- Kentucky Fried Chicken, 534, 536
- Kenya, 659
- Kerin, Paul, 468
- Killinger, Kerry, 717
- King's Dominion, 316
- kinked demand curve, 439–40, 441
- Kloppers, Marius, 237, 468–69
- Knott's Berry Farm, 316
- Köhler, Horst, 596
- Kravetz, Shawn, 16
- Krueger, Alan, 323
- Kulovaara, Harri, 191
- Kuwait, 421
- Kyoto Protocol, 359

- labor productivity, *see* productivity
- Lagarde, Christine, 596
- lagging series, 806
- Lagrangian multipliers, 168–71, 754–56, 758
- Las Vegas Sands Corporation, 600
- Latham, Scott, 588, 589
- law of diminishing marginal returns, 143
- leading indicators, 806–7, 812
- leading series, 806–7
- learning curve, 777–80, 786
- least squares, method of, 106–8, 106n3, 124
- legal constraints, 5
- “lemons” in used car market, 619, 620–22
- leverage, 604
- Libya, 420
- life insurance, adverse selection and, 631–33
- lignin, 149
- linear trends, estimating, 795–97
- LinkedIn, 149
- Lipitor, 534
- long-run adjustment process, 247–51, 252
- long-run cost functions, 184–89, 204
 - average, 184–89
 - economies of scale and, 189–91
 - economies of scope and, 191–94, 204
 - marginal, 188–89
 - total, 188–89
- long-run equilibrium, 245–47, 248, 250, 252, 284–85
- Long's Drug Stores, 46, 335
- Lopez, Jose Ignacio, 424n5
- Lucas, George, 4
- Lynch, Mike, 145

- Mahindra, Anand, 513
- Mama Cares Foundation, 682
- managerial economics
 - CEO pay and, 586–87
 - microeconomics vs., 2–3, 20
- managerial examples
 - of adverse selection, 624, 626, 628, 630, 636–37
 - of antitrust violations, 678
 - of auctions, 503, 512–13, 516–17
 - of bundling, 359, 361, 383
 - of call options, 594
 - of cartels, 420
 - of cause marketing, 4
 - of coefficient of determination, 111
 - of corporate synergy in conglomerates, 194
 - of cost-plus pricing, 268
 - of duopoly, 418, 422, 430, 432
 - of economies in oil tankers, 165
 - of economies of scope, 192–93
 - of elasticity, 46–47, 49, 51
 - of environmental pollution regulation, 710–11
 - of estimating demand, 54
 - of forecasting, 239
 - of franchiser vs. franchisee, 266
 - of futures markets, 323
 - of game theory, 468–69
 - of government-business relations, 659, 662, 666, 682–83, 715–19
 - of innovations in payments for goods and services, 152
 - of marginal cost, 237
 - of marketing plans, 102
 - of market movement, 18
 - of monopolies, 262, 270, 662, 682–83
 - of moral hazard, 608, 611, 636–37

- of oligopoly, 414
- of output and profit trade-off, 77
- of parallel development efforts, 773
- of price discrimination, 307, 313, 316, 325, 328, 334
- of price elasticity of demand, 275
- of principal-agent issues, 575, 586–89, 593, 596–98, 600–601, 604–5
- of production function, 137, 145, 147
- of risk, 534, 535, 544–45, 555, 556
- of risk and return trade-off, 91
- of scientific pricing for great art, 335
- of servicing demand in emerging markets, 34
- of sophisticated pricing, 326, 516–17
- of substitution, 149, 155
- of supply and demand, 8, 10, 12, 16
- of time constraint on consumer behavior, 75
- of transfer pricing, 394
- of two-part tariffs, 329, 331
- of unbundling, 384
- of value of time and full price of transportation, 218–19
- of whistle-blowing, 688–89
- of winners curse, 521–22
- of yield management and performance, 317
- managers
 - flat salary and, 573–74, 584
 - incentive compatibility and, 579–81
 - incentive compensation and, 575–79, 583–84, 600–601
 - main task of, 2, 20
 - objectives of, 3
 - profit sharing and, 585, 591
 - stock options and, 591–94, 604–5
 - see also* principal-agent issues
- Mannesman, 523
- Marathon Oil, 193
- marginal analysis, 729–31
- marginal cost, 178, 179–82, 184
 - duopolies and, 418–19, 419n3, 425–26
 - monopolies and, 261, 262–66, 270, 289
 - optimization and, calculus of, 748–50
 - perfect competition and, 232–38, 240–42
 - of pollution and pollution control, 707
 - rising, two-part tariff with, 330, 333, 340
- marginal cost function, long-run, 188–89
- marginal expenditure, 241
- marginal product, 141, 142, 143, 181
- marginal profit, 729–30, 731–34
- marginal rate of substitution, 68–69, 76, 93
- marginal rate of technical substitution (MRTS), 149–52, 165
- marginal revenue
 - duopolies and, 421–23, 426, 429
 - monopolies and, 261–66, 273, 289, 290
 - optimization and, calculus of, 748–50
 - perfect competition and, 231–34, 235–36
 - price elasticity and, 48–53, 61
 - total, 276–77
- marginal revenue product (MRP), 240–41
- marginal values, 729, 730–34, 735, 758
- market bundles
 - equilibrium, 73–75, 76, 85
 - indifference curves and, 66–68, 92–93
 - utilities of, 69–70
- market-clearing price, *see* price
- market demand curve, 11–13, 21, 29–32, 60
 - consumer income levels and, 30–31, 32
 - consumer tastes and, 30, 32
 - for copper, 11
 - demand function and, 33–35
 - for food in 2008, 18–19
 - individual demand curves and, 88–90
 - monopolies and, 257–58, 259–61, 264
 - in perfect competition, 225, 227–29
 - population and, 31
 - shifts in, 17–18, 21, 34, 60, 229
 - with zero and infinite price elasticities of demand, 37
- market demand function, 60–61
 - definition of, 33
 - demand curve and, 33–35
 - parameters of, 33
 - in price elasticity of demand, 40–41
- market demand schedule, 29, 39
- market experiments, 101, 102, 123–24
- market four-firm concentration ratio, 668–74
- market leaders, 417
- market price in perfect competition, 225, 227–30
- markets
 - definition of, 10
 - emerging, servicing demands in, 34
 - movement of, life during, 18–19
 - segmenting and sealing of, 196, 312, 312n2
 - see also* demand and supply, principles of
- market structures, 224–27
- market supply curve, 13, 18–20, 21
 - for food in 2008, 18–19
 - in perfect competition, 225, 227–29
 - shifts in, 18–20, 21, 227–29
- markup pricing, 326
- markups, 271–73, 290
- Massachusetts Department of Public Utilities, 666
- matching games, 490
- matrix form, 464, 469–72
- maximization and minimization problems,
 - derivatives to solve, 744–48, 758
- MBS (mortgage-backed securities), 716, 717–19
- McDonald's, 4, 536
 - franchisee requirements of, 676
 - tying by, 388
- Medicare, 630
- Merck, 193, 534
- method of least squares, 106–8, 106n3, 124
- Mexico, 683, 685, 686
- microeconomics vs. managerial economics, 2–3, 20
- Microsoft
 - antitrust trial of, 656
 - tying by, 386–88
- Middle East, oil and, 421–22
- minimills, 147
- minimization and maximization problems,
 - derivatives to solve, 744–48, 758
- Minitab, 113–14
- mixed bundling, 357–77
- Model T automobile, 779–80
- Monitoring Analytics, 237
- monopolistic competition, 226–27, 257–59, 282–85, 290
 - advertising expenditures in, 285–89, 290
 - definition of, 226
 - price and output decisions in, 283–85, 298–301
- monopoly, 225, 226, 227, 289
 - cartels and, 420–21
 - competition vs., 658–60
 - cost-plus pricing and, 267–73, 289
 - Excel exercise on, 294–98

INDEX

- monopoly (*continued*)
 - government regulation of, 660–64
 - joint product pricing and, 274, 276–81, 290
 - monopsony and, 281–82, 290
 - multiple-product firms and, 273–74
 - OPEC and, 420
 - price and output decisions in, 259–65, 267, 283–85, 289, 298–301
 - problems on, 267, 273, 275, 278, 290–94
 - regulated, efficiency of, 666–67
 - social cost of, 662
 - social issues and, 682–83
- monopsony, 281–82, 290
- Montana-Dakota Utilities Co., 664
- Moody's, 418, 432, 600
- moral hazard, 716, 717
 - adverse selection and, 636–37, 644–47
 - definition of, 594
 - Excel exercise on, 616–18
 - financial crisis of 2008 and, 604, 611
 - health care debate and, 608
 - insurance and, 594–95
 - see also* principal-agent issues
- mortality rates
 - in United Kingdom, 633
 - in United States, 631–33
- mortgage-backed securities (MBS), 716, 717–19
- MRP, *see* marginal revenue product (MRP)
- MRTS, *see* marginal rate of technical substitution (MRTS)
- Muhlemann, Max, 328
- multicollinearity, 118–19, 125
- multiequation models, 807–9
- multinational firms, transfer pricing and, 394
- multiple coefficient of determination, 114, 114n7
- multiple regression, 109, 112–14, 113n6, 123, 124, 125
- multiple sclerosis, 49
- multi-product firms, pricing by, 273–74
- multivariable functions, partial differentiation and maximization of, 750–52
- NAFTA (North American Free Trade Agreement), 696
- Nash, John, 427, 472
- Nash equilibrium, 427, 472–78, 483, 488, 489, 509
- national parks, entrance fees to, 715
- National Park Service, 715
- National Research Council, 183
- near-field communication (NFC), 152
- negative correlation, 357
- negative serial correlation, 120
- negotiation costs, 195
- Netflix, 384, 478
- Netherlands, 686, 689–91, 693–94, 695
- Netscape Navigator, 387–88
- network economies, 198–201
- networks in game theory, 477–78
- New Yorker*, 383
- New York Jets, 503
- New York Yankees, 137
- New Zealand, 664, 710–11
- NFC (near-field communication), 152
- nonlinear trends, estimating, 797–98
- normal goods, 53
- North American Free Trade Agreement (NAFTA), 696
- Northeast Utilities, 710
- Northrop Grumman, 720
- Novartis AG, 49
- Nucor, 147
- numerically controlled machine tools, forecasting
 - diffusion rate of, 782–85
- nursing homes, economies of scale in, 189–90
- Nutriset, 682, 683
- Oasis of the Seas*, 191
- Ocean Spray, 414
- Office Depot, 679
- oil tankers, economies in, 164
- oligopoly, 226, 227, 410–59
 - collusive agreements in, 413–14, 440
 - cooperative behavior and, 411–13, 440
 - cranberry production as, 414
 - definition of, 410
 - price leadership in, 414–16, 440, 447–49
 - problems on, 416, 441–46
 - sticky prices in, 439–40, 441
 - see also* duopoly
- OPEC (Organization of Petroleum Exporting Countries), 420–21, 686
- operating leverage, break-even analysis and, 211–14
- opportunity cost doctrine, 173
- opportunity costs, 173–74, 200, 201, 204
- optimization techniques, 728–61
 - constrained optimization, 752–54, 758
 - derivatives, 734–37, 758
 - finding, 738–44
 - to solve maximization and minimization problems, 744–48, 758
 - functional relationships, 728–29, 758
 - incremental costs and incremental revenues, comparing, 756–58
 - Lagrangian multipliers, 754–56, 758
 - marginal analysis, 729–31
 - marginal cost equals marginal cost rule, 748–50
 - partial differentiation and maximization of multivariable functions, 750–52
 - problems on, 758–61
 - relationships among total, marginal, and average values, 731–34
- Organization of Petroleum Exporting Countries (OPEC), 420–21, 686
- Orszag, Peter, 215
- output
 - allocation of, among plants, 298–301
 - costs and, 183
 - of joint products, 277–81
 - in monopolistic competition, 283–85
 - under monopoly, 259–65, 267, 289
 - in perfect competition, 229–34, 235, 236–38, 240–42, 245, 252
 - per worker, 140
 - regulation-induced cost increases and, 710–14
 - of statistical software, interpreting, 114–18, 115n8, 118n9
- output elasticity, 160–61
- own-price elasticity of demand, *see* price elasticity of demand
- Pabst, 678
- Palm Management Corporation, 275
- parallel development efforts, 770–71, 773
- parcel tanker business, 422–23
- Parker Hannifin, 326
- partial differentiation and maximization of multivariable functions, 750–52
- patents, 680–81, 684, 720

- Paulson, Hank, 716
- PDP (prescription drug plan), 630
- peak load pricing, 319–22, 324
- peaks in business fluctuations, 802, 807
- Pepsi, 522
- Pepsico, 555
- perfect competition, 224–56
 - definition of, 226
 - Excel exercise on, 254–56
 - long-run adjustment process in, 247–51, 252
 - long-run equilibrium in, 245–47, 248, 250, 252
 - market price in, 225, 227–30
 - market structures and, 224–27
 - output decisions in, 229–34, 235, 236–38, 240–42, 245, 252
 - problems on, 239, 245, 252–54
 - resource allocation in, 251
 - short-run producer surplus in, 242–45, 252
 - transfer pricing and, 392–93, 395
- perfect information
 - adverse selection and, 622–23
 - expected value of, 540–42, 558
- per-unit taxes, 702–3
- Pfizer, 534
- Philadelphia Eagles, 328
- philantreneurs, 4
- Philip Morris, 43
- Philippines, 433
- PJM Interconnection, 237
- plants, allocation of output among, 298–301
- Plumpy' nut, 682–83
- point elasticity of demand, 39–40
- pollution, *see* environmental pollution, regulation of
- population, market demand curve and, 31
- population regression line, 104
- positive serial correlation, 119–20, 119n10
- power functions, derivatives of, 739–40
- prescription drug plan (PDP), 630
- present value, 3–5, 821–27
- price
 - actual, 15–17
 - equilibrium, 14–15, 16–17, 18–20, 21, 228–29
 - marginal cost and, in perfectly competitive firm, 234–36, 238, 240–42
 - market, in perfect competition, 225, 227–29
 - regulation-induced cost increases and, 710–14
- price ceilings, 698–702
- price discrimination, 92, 304–56
 - airlines and, 304, 316–18
 - coupons and, 318–19
 - definition of, 306
 - Excel exercise on, 344–46
 - first-degree, 196, 306–11, 329, 339
 - motivation for, 305–6
 - problems on, 308–9, 320–21, 332, 336–37, 340–43
 - rebates and, 318–19
 - second-degree, 311–12, 339
 - third-degree, 312–16, 318, 320–21, 322, 339, 346–51
 - two-part tariffs and, 324–26, 327, 329–39
- price elasticity of demand, 28–29, 36–53, 61
 - advertising and, 57–59, 288–89, 290
 - arc elasticity and, 39–40
 - constant-elasticity demand function and, 59–60, 61
 - coupons and, 318–19
 - cross, 55–57, 58, 61
 - definition of, 36
 - demand function in, 40–41
 - determinants of, 44–45
 - examples of, 43, 45–48
 - infinite, 36, 37
 - markup and, 271–73
 - monopolies and, 262, 290
 - point elasticity and, 39–40
 - revenues and, 41–43, 48–53, 61
 - specialty drugs and, 49
 - strategic use of, 45–48
 - unitary, 36, 37, 59–60
 - zero, 36–37
- price floors, 698–702, 720
- price leadership, 414–16, 440, 447–49
- Priceline.com, 516–17
- pricing
 - congestion, 218
 - cost-plus, *see* cost-plus pricing
 - dynamic, 9–10
 - of joint products, 274, 276–77
 - in monopolistic competition, 283–85
 - in monopoly, 259–65, 267, 275
 - by multi-product firms, 273–74
 - peak load, 319–22, 324
 - transfer, *see* transfer pricing
 - tying, 386–88
 - value-based, 46
 - variable, 8
- principal-agent issues, 6–7, 21, 567–618
 - asset substitution in, 595, 599, 602–3, 606, 607
 - conflicts between owners and managers, 569–70
 - examples of, 570–72
 - incentive pay, 600–601
 - managerial behavior and effort in, 572–74
 - nonobservable effort and incentive conflict in, 579–81
 - observable effort and incentive conflict in, 574–79
 - private equity firms and, 575
 - problems on, 578–79, 583, 584–85, 613–16
 - product liability law and, 607–12
 - risky output and nonobservable effort in, 578–79, 581–90
 - between shareholders and creditors, 606–7
 - see also* managers
- printers, expense of, 387
- prisoner's dilemma, 483–85, 488
- private equity firms, 575
- private-value auctions, 506, 523
- probability and risk, 533–34, 536, 557
- probability distributions, 536–37, 551–53
- producer surplus, 242–45, 252, 691–94, 695
- product differentiation
 - in duopolies, 435–38
 - in monopolistic competition, 283
- product groups, 283
- production theory, 136–71
 - estimations of, 161, 164–65
 - Excel exercise on, 209–10
 - innovation and, 149
 - isoquants in, 146, 148, 150, 151, 162–63, 165
 - law of diminishing marginal returns and, 143
 - marginal rate of technical substitution and, 149–52
 - with one variable input, 137–43
 - optimal input combination in, 152–56, 168–71
 - output elasticity and, 160–61
 - problems on, 158–59, 162–63, 165–68
 - returns to scale and, 156–60
 - with two variable inputs, 143–44, 144n1, 146

INDEX

- productivity, 765–66
 - information technology and, 784
 - total factor, 766–68, 785
- product liability law, consumer goods safety and, 607–12
- profit
 - accounting vs. economic concept of, 5–6, 20–21
 - average, 730
 - definition of, 5–6
 - expected value of, 537–38
 - innovation and, 6, 21
 - marginal, 729–30
 - marginal cost equals marginal cost rule for, 748–50
 - monopolies and, 259–61
 - reasons for, 6, 21
 - risk and, 6, 21
 - see also* profit maximization
- profit contribution analysis, 203–4
- profit margin, definition of, 268
- profit maximization, 748–49
 - advertising expenditures and, 285–88
 - in auctions, 510–12
 - cost-plus pricing and, 271–73, 289
 - duopolies and, 426–27, 441
 - joint product pricing and, 278–79
 - single-price monopolist, 305–7
 - two-part tariffs and, 327, 329
- profit sharing, manager motivation through, 585, 591–94
- public goods, 714, 718–20, 721
- Public Service Enterprise Group, 710
- public transit, price elasticity and, 43
- public utilities, regulation of, 664–66, 720
- Public Utilities Commission (PUC), 660–64
- PUC, *see* Public Utilities Commission (PUC)
- Puget Sound Energy, 325
- Putnam, Howard, 675
- Qatar, 421
- quotas, 694–95
- quotients, derivatives of, 743–44
- railroads, 54, 197, 270
- rate of return, internal, 826–27
- rate-of-return model, 660–64
- rational choice, 93
- reaction function, 427–28
- Reagan, Ronald, 111
- rebates in price discrimination, 318–19
- recession, 54, 802
- Red campaign, 4
- regression analysis, 101–3, 113, 124–25
 - computer printouts of, 113–14, 114n7, 115
 - multicollinearity and, 118–19
- regression coefficient, estimated, 107, 124
- regression line, sample, 105–6, 108
- regression models, simple, 103–4, 109
- regulatory lags, 667
- Reilly, John, 111
- repeated games, 483–85, 495
- repurchase tender offers (RTOs), 515–18
- reputation building, 487, 489, 570
- research and development, 769–72, 786
 - parallel development efforts, 770–71, 773
 - project selection in, 772–74
- reservation price
 - in auctions, 196, 504, 510, 523, 524
 - consumer, 90
- Reserve Mining Company, 703, 710
- residual claims, 599
- residuals, 120–21, 123, 810–12
- resource allocation in perfect competition, 251
- returns to scale, 156–60
- return vs. risk, 91
- revenue equivalence theorem, 507
- revenues
 - incremental, 756–58
 - marginal, *see* marginal revenue
 - price elasticity and, 41–43, 48–53
- reverse auctions, 512–13, 516–17
- ridge lines, definition of, 152
- risk, 532–66
 - attitudes toward, 542–43, 546–50
 - certainty equivalence, insurance, and, 554–57
 - decision trees and, 538–41, 557
 - definition of, 533
 - discount rates, risk-adjusted, 556
 - Excel exercise on, 562–66
 - expected profit and, 537–38, 557
 - expected values and, 537
 - measures of, 551–53, 558
 - perfect information and, 540–42
 - probability and, 533–34, 536, 557
 - probability distributions and, 536–37
 - problems on, 540–41, 543, 547, 558–62
 - profit and, 6, 21
 - return vs., 91
 - valuation model for, 553–54, 558
- risk averters, 518, 548, 549n5, 550
- risk lovers, 549, 549n4, 550
- risk-neutral, 549, 550
- risk premiums, 582
- risk sharing vs. efficiency in executive compensation, 581–89
- Robinson, Larry, 46
- Robinson-Patman Act (1936), 676, 720
- Rockefeller, John D., 411
- Rodriguez, Alex, 137
- Rommel, Rick, 624
- Ronald McDonald houses, 4
- Rothschild, Michael, 634
- Royal Caribbean, 191
- royalties, *see* total revenue
- RTOs, *see* repurchase tender offers (RTOs)
- rule of reason, 677
- Russia, 359
- Safaricom, 659
- salmon, forecasting the price of, 239
- sample regression line, 105–6, 108
- Samuelson, Paul A., 810–11
- Sanderson, Allen, 715
- Sarkozy, Nicolas, 596
- SAS printouts, 114, 115
- Saudi Arabia, 421
- scale economies, 189–91, 411
- scale of plant, 174
- scientific pricing for great art, 335
- scope, economies of, 3, 191–94, 204
- Seachem, Odfjell, 423
- sealed-bid auctions, 505, 507, 508
- sealing of markets, segmenting and, 196, 312, 312n2
- Sears auto repair, 600
- seasonal indexes, 793, 799, 812

- seasonal variation, 798–802, 812
- seat licenses for sports teams, 328
- second-degree price discrimination, 311–12, 339
- second-price auctions, 505, 508, 523
- Securities and Exchange Commission, 657
- segmenting and sealing of markets, 196, 312, 312n2
- self-selection, resolving adverse selection through, 634–38
- self-selection menus, 638
- separating equilibrium, 638
- serial correlation, 119–20, 119n10
- Sherman Act, 674–75, 676, 720
- shipping, cartels in, 420
- Shopkick, 624
- short-run cost functions, 174–77, 185, 187, 204
 - average fixed cost, 177, 183
 - average total cost, 179
 - average variable cost, 177–79, 181–82
 - mathematical form for, 214–17, 220–21
 - total costs, 176
 - total fixed costs, 175–77, 183
 - total variable costs, 175–76
- shutdown point, 238
- simple bundling, 357–58
 - see also* bundling
- simple regression models, 103–4, 109, 124
- simultaneous behavior, 417
- Singapore, 686
- Skype, 478
- Smith, Adam, 17, 225
- Smith, Fred, 475
- smoothing constant, 818
- sniping, 505
- social welfare triangle, 662
- solarcoaster, 16
- solution concepts in game theory, 467
- sophisticated pricing, 196, 326
- specialty drugs, elasticity in, 49
- Spence, Michael, 639
- spinoffs, 193
- sports teams, seat licenses for, 328
- Stackelberg behavior, 431–35, 441, 457–59
- standard deviation, 551–53, 551n7, 558
- standard error of estimate, 114–19, 115n8, 124–25
- Staples, 679
- statistical software, interpreting output of, 114–18, 115n8, 118n9
- steel industry, 147
- Stemberg, Tom, 679
- sticky prices, 439–40, 441
- Stiglitz, Joseph, 634
- stockholders, 7
- stock investing, 16
- stock options, manager motivation and, 591–94, 604–5
- Stolt-Nielsen SA, 422–23
- strategic foresight, 479–83, 495
- strategy in game theory, 460–63, 494
- strategy session examples, *see* managerial examples
- strictly competitive games, 493–94
- strike price, 591–92, 594
- subgame, 483
- subjective definition of probability, 536
- substitutes and price elasticity of demand, 44–45
- sums and differences, derivatives of, 740–43
- sunk costs, 174
- Suntech, 16
- Sun Tzu, 462
- Super Bowl tickets, 323
- supply and demand, *see* demand and supply, principles of
- supply curve, market, *see* market supply curve
- Surface Transportation Board, U.S., 270
- survey techniques, 792–93, 812
- sustainable market price, 15
- Tankers, Jo, 423
- target return, 268
- tariffs, 694–95, 720
 - two-part, 324–26, 327, 329–39
- TARP (Troubled Assets Relief Program), 716, 718, 719
- taxes
 - transfer pricing and, 394, 395–98
 - welfare impacts of, 702–3, 721
- technological change, 763–89
 - definition of, 763, 785
 - diffusion models and, 780–85
 - learning curve and, 777–80, 786
 - problems on, 784, 786–89
 - rate of, 765–68
 - research, development, and, 769–74, 786
 - see also* innovation
- Texas Instruments, 777, 779
- textbooks, bundling, 361
- Tharp, Charles, 596
- theory of the firm, 3–5, 20
- Therma-Stent, 269
- Thieberger, Victoria, 410, 411
- third-degree price discrimination, 312–16, 318, 320–21, 322, 339, 346–51
- Thompson, G. Kennedy, 605
- time constraints, consumer behavior and, 75
- time-cost trade offs, 774–77, 786
- time series, 793–95, 802, 812
- Times of India, The*, 91
- tit for tat in games, 486
- T-Mobile, 523
- total costs
 - long-run, 188–89
 - monopolies and, 259–62
 - perfect competition and, 231, 232
 - short-run, 176
- total factor productivity, 766–68, 785
- total fixed costs, short-run, 175–77, 183
- total marginal revenue curve, 276–77
- total revenue
 - duopolies and, 419, 419n3, 429
 - monopolies and, 259–61, 266, 273–74
 - price elasticity and, 41–43, 48–53
- total values, 731–34
- total variable costs, short-run, 175–76
- Toyota Camrys, 621
- Toyota Motor Company, 5, 621
- trade and trade policy, 682–83, 685–98, 720
 - demand and supply in, 688–91
 - foreign, 682–83, 685–88
 - free, producer and consumer surplus and, 691–94, 695
 - strategic, 696–98
 - tariffs and quotas in mitigating gains from, 694–95, 720
- transaction costs, 194–97
- transferable emissions permits, 708

INDEX

- transfer pricing, 388–99
 - definition of, 388
 - Excel exercise on, 403–7
 - global use of, 395–98
 - perfect competition and, 392–93, 395
- Transnet Freight Rail, 270
- transportation
 - deregulation of, 656
 - full cost of, 197, 218–19
- travel time, value of, 218–19
- trend values, definition of, 796
- Trenton Gas Company, 665
- Troubled Assets Relief Program (TARP), 716, 718, 719
- trough, 802, 807
- true regression line, 104
- t statistic, 117–18, 118n9, 125
- Tupperware, 34
- Turner, Ted, 4
- two-part tariffs, 324–26, 327, 329–39
 - definition of, 324
 - with different demand curves, 331, 333–39
 - with intersecting demands, 351–56
 - with rising marginal costs, 330, 333, 340
- tying, 386–88, 676
- UBid, 504
- unbundling, 377–81, 384–85
- unitary elasticity, 36, 37, 59–60
- United Technologies, 193
- University of California at Berkeley, 334
- University of Kansas, 334
- UPS, 475–77
- Urban Institute, 626
- used automobiles, 619, 620–22, 648
- utility, 69–70
 - definition of, 69
 - maximizing of, 76, 77, 79–84
- utility function, 546
 - construction of, 546–48, 558
 - types of, 548–51
- utility pricing, 311–12
- value
 - expected, 537, 540–42, 543, 800
 - marginal, 729, 730–34, 758
 - risk and, 553–54
- value-based pricing, 46
- value of firm, 3–5, 20
- variable costs
 - average, 177–79, 181–82
 - total, 175–77
- variable inputs, 174–77
 - one, production function with, 137–43
 - two, production function with, 143–44, 144n1, 146
- variable pricing, 8
- Venkatasubramanian, Venkat, 596
- venture capitalists, 636–37, 644–47
- Verified Identity Pass, 218
- Verizon, 97, 331
- Vickrey auctions, *see* second-price auctions
- Vivendi, 469
- Volkswagen, 424n5
- W. W. Norton, 193
- Waldfoegel, Joel, 84
- Wal-Mart, 327, 414–15, 784
- Walt Disney Company, 503
- WaMu (Washington Mutual Savings), 717
- War of 1812, 502
- warranties as signals, 640–43
- Washington Mutual Savings (WaMu), 717
- Washkewicz, Donald, 326
- water, demand for, 214–15
- welfare impacts of taxes, 702–3, 721
- Wellcome, 523
- Wesfarmers Ltd., 410
- Wharton model, 809
- Wharton School, 173
- Wheelwright, Steven, 767
- whistle-blowing, 688–89
- Whitaker, Pulina, 596
- White, Matt, 378n7
- Wilcox, John, 597
- Williams, Richard, 97
- willingness to pay (WTP), 90
- winner's curse, 520–22, 524
- wireless phone industry, 325
- Wolff, Patricia, 682
- Woolworth Ltd., 410
- World Economic Forum, 535
- Wyoming, 664
- Xerox Corporation, tying by, 386–88
- Xingdong, Fang, 679
- Yellowstone National Park, 715
- yield management
 - airlines and, 45–46, 317
 - sports teams and, 10
- zero-sum games, 494
- Zuckerberg, Mark, 4

Real-world case studies—the Strategy Sessions—in the Eighth Edition of *Managerial Economics*:

1. Bono See Red, and Corporations Profits See Black – 2008, p. 4
2. Baseball Discovers the Law of Supply and Demand – NEW, 2011, pp. 8–10
3. Supply and Demand in Action – NEW, 2011, p. 12
4. Using the Law of Supply and Demand to Invest in Stocks – NEW, 2011, p. 16
5. Life During a Market Movement – 2008, pp. 18–19
6. The Customer Is Always Right—Wrong! – 2004, p. 31
7. Servicing Demand in Emerging Markets – NEW, 2011, p. 34
8. Elasticity in Use – 2002, p. 46–47
9. Elasticity in Specialty Drugs – NEW, 2011, p. 49
10. Beer—Elastic or Inelastic? – NEW, 2011, p. 51
11. Estimating the Demand for Amtrak Rail Passenger Business – 2008, p. 54
12. The Effect of a Time Constraint on Consumer Business – 2008, p. 75
13. A Manager's Trade-Off Between Output and Profit – 2008, p. 77
14. The Trade-Off between Risk and Return – 2008, p. 91
15. Marketing Plans at the Stafford Company – NEW, 2011, p. 102
16. How the Japanese Motorcycle Makers Used the Coefficient of Determination – NEW, 2011, p. 111
17. The Yankees' Deal for Alex Rodriguez – 2004, p. 137
18. Substitution of Capital for Labor in Legal Proceedings – NEW, 2011, p. 145
19. How Nucor Stays on the Production Function – 2008, p. 147
20. Substitution – NEW, 2010, p. 149
21. Innovations in Payments for Goods and Services – NEW, 2011, p. 152
22. The Substitution of Inputs in Baseball – NEW, 2010, p. 155
23. Economies in Oil Tankers – 2008, p. 164
24. Economies of Scope in Advertising Agencies – 2008, p. 192
25. Economies of Scope – NEW, 2011, p. 193
26. Corporate Synergy in Conglomerates – 2007, p. 194
27. Mr. Martin Gets Chewed Out by the Boss – 2008, p. 203
28. Water, Water, Not Everywhere – NEW, 2011, pp. 214–215
29. The Value of Time and the Full Price of Transportation – 2007, pp. 218–219
30. Competition and Price Equals Marginal Cost – NEW, 2011, p. 237
31. Forecasting the Price of Salmon – 2008, p. 239
32. Why Monopoly Power Attracts Warren Buffet – NEW, 2011, p. 262
33. Franchiser versus Franchisee? – 2008, p. 266
34. Markup Pricing on the iPad2 – NEW, 2011, p. 268
35. When Is the Price Too High? – NEW, 2011, p. 270
36. When Can You Haggle? – 2008, p. 307
37. That Darling Little Mouse Is Really a Price Discriminator – 2007, p. 313
38. Mickey Mouse Pricing at Amusement Parts – 2004, p. 316
39. Yield Management and Airline Performance – 2002, p. 317
40. The Future Is Now: The Futures Market for Super Bowl Tickets – 2006, p. 323
41. Why Do Your Laundry at 3 a.m.? – 2002, p. 325
42. A Change from Mark-Up Pricing to Sophisticated Pricing – 2007, p. 326
43. Making Them Pay Twice: Personal Seat Licenses for Sports Teams – 2008, p. 328
44. Costco and the Two-Part Tariff – 2007, p. 329
45. Verizon Local Calling Plans – 2008, p. 331
46. Academic Institutions Practice What They Preach – NEW, 2010, p. 334
47. Scientific Pricing—Even for Great Art? – 2006, p. 335
48. Bundling Carbon Credits with Gas Sales – 2007, p. 359
49. Bundling College Textbooks – NEW, 2010, p. 361
50. How *The New Yorker* Used Bundling – 1999, p. 383
51. When to Unbundle – NEW, 2011, pp. 384–385

52. Which Is More Expensive? Printers or Ink? – 2007, p. 387
53. Transfer Pricing in Practice – NEW, 2011, p. 394
54. Cranberries: Where 30% of the Market Are Price Takers – 2007, p. 414
55. Warren Buffett and Duopoly – NEW, 2011, p. 418
56. Cartels Come in Many Shapes and Sizes – 1997, pp. 420–421
57. How Seagoing Chemical Haulers May Have Tried to Share the Market – 2004, pp. 422–423
58. The End of a Duopoly? – NEW, 2011, p. 430
59. Competition among the Few – NEW, 2011, pp. 432–433
60. Managerial Use of Game Theory – 2003, pp. 468–469
61. The Expanding Use of Auctions – NEW, 2011, p. 503
62. How Managers Use Reverse Auctions to Reduce Cost – 2008, pp. 512–513
63. The Use of Sophisticated Pricing within an Auction Format – 2008, pp. 516–517
64. The Winner's Curse in Bidding for Oil Rights – 2008, p. 521
65. Why Managers Must Care About Profit, Not Market Share – NEW, 2011, p. 522
66. Pfizer and Its New Cholesterol Drug – 2006, p. 534
67. What Are the Major Global Risks in 2011? – 2011, p. 535
68. FedEx's and Ata Holdings Corporation's Disclosures about Risk – 2007, pp. 544–545
69. Pepsico Risk Management – 2007, p. 555
70. The Use of Risk-Adjusted Discount Rates – 2008, p. 556
71. Private Equity Firms Deal with the Principal-Agent Problem – NEW, 2011, p. 575
72. The Value of CEO Pay after the 2008 Crash – NEW, 2011, pp. 586–587
73. Principal-Agent Issues in Failed IPOs – NEW, 2011, pp. 588–589
74. Getting the Board to Focus on the Long Term – 2006, p. 593
75. Call Options – 2008, p. 594
76. Executive Pay on the Continent – NEW, 2011, pp. 596–598
77. The Good and Bad of Incentive Pay – 2006, pp. 600–601
78. The Song of the Sirens – 2008, pp. 604–605
79. Moral Hazard and the National Health Care Debate – NEW, 2010, p. 608
80. Moral Hazard in the Financial Market – 2008, p. 611
81. What Does Big Brother Know and When Does He Know It? – NEW, 2010, p. 624
82. Adverse Selection and the National Health Care Debate – NEW, 2010, p. 626
83. Information Asymmetries and Brownfields – 2005, p. 628
84. Adverse Selection in the Federal Government Prescription Drug Plan – 2006, p. 630
85. How Do Venture Capitalists Resolve Adverse Selection and Moral Hazard Problems? – NEW, 2011, pp. 636–637
86. Government Actions and the Financial Crisis of 2008 – 2009, p. 657
87. More Competition in Kenya – NEW, 2010, p. 659
88. The Social Cost of Monopoly – 2008, p. 662
89. A Dispute over a Requested Gas Rate Increase – 2008, p. 666
90. The Art of the Deal Is Back – NEW, 2011, p. 674
91. Antitrust Violations – NEW, 2009, p. 678
92. China: New to Antitrust – NEW, 2011, p. 679
93. Antitrust on the Continent – NEW, 2011, p. 680
94. Monopoly and Social Issues – NEW, 2011, pp. 682–683
95. Using Patents to Maintain Market Power – NEW, 2011, p. 684
96. Making Whistle-Blowing Pay Off – 2008, pp. 688–689
97. Buying and Selling the Right to Emit Greenhouse Gases – NEW, 2011, pp. 710–711
98. Entrance Fees to National Parks – 1995, p. 715
99. All the Glitters is Not Gold – 2008, pp. 716–719
100. Parallel Development Efforts at IBM – 1988, p. 773