WILEY FINANCE

MODELING RISK

+ CD

Applying Monte Carlo Simulation, Real Options Analysis, Forecasting, and Optimization Techniques

The CD-ROM includes trial versions of Risk Simulator and Real Options Super Lattice Solver software. Also included are sample training videos.

JOHNATHAN MUN

Modeling Risk

Founded in 1807, John Wiley & Sons is the oldest independent publishing company in the United States. With offices in North America, Europe, Australia, and Asia, Wiley is globally committed to developing and marketing print and electronic products and services for our customers' professional and personal knowledge and understanding.

The Wiley Finance series contains books written specifically for finance and investment professionals as well as sophisticated individual investors and their financial advisors. Book topics range from portfolio management to e-commerce, risk management, financial engineering, valuation, and financial instrument analysis, as well as much more.

For a list of available titles, visit our Web site at www.WileyFinance.com.

Modeling Risk

Applying Monte Carlo Simulation, Real Options Analysis, Forecasting, and Optimization Techniques

JOHNATHAN MUN

WILEY

John Wiley & Sons, Inc.

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data:

Mun, Johnathan.
Modeling risk : applying Monte Carlo simulation, real options analysis, forecasting, and optimization techniques / Johnathan Mun
p. cm.—(Wiley finance series)
Includes index.
ISBN-13 978-0-471-78900-0 (cloth/cd-rom)
ISBN-10 0-471-78900-3 (cloth/cd-rom)
1. Risk assessment. 2. Risk assessment—Mathematical models. 3. Risk management. 4. Finance—Decision making. I. Title. II. Series.

HD61.M7942 2006
658.15—dc22 2005058120

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

To my wife Penny, the love of my life.
In a world where risk and uncertainty abound,
you are the only constant in my life.

*Delight yourself in the Lord and
he will give you the desires of
your heart.*

Psalms 37:4

Preface

We live in an environment fraught with risk and operate our businesses in a risky world, as higher rewards only come with risks. Ignoring the element of risk when corporate strategy is being framed and when tactical projects are being implemented would be unimaginable. In addressing the issue of risk, *Modeling Risk* provides a novel view of evaluating business decisions, projects, and strategies by taking into consideration a unified strategic portfolio analytical process. This book provides a qualitative and quantitative description of risk, as well as introductions to the methods used in identifying, quantifying, applying, predicting, valuing, hedging, diversifying, and managing risk through rigorous examples of the methods' applicability in the decision-making process.

Pragmatic applications are emphasized in order to demystify the many elements inherent in risk analysis. A black box will remain a black box if no one can understand the concepts despite its power and applicability. It is only when the black box becomes transparent so that analysts can understand, apply, and convince others of its results, value-add, and applicability, that the approach will receive widespread influence. The demystification of risk analysis is achieved by presenting step-by-step applications and multiple business cases, as well as discussing real-life applications.

This book is targeted at both the uninitiated professional and those well versed in risk analysis—there is something for everyone. It is also appropriate for use at the second-year M.B.A. level or as an introductory Ph.D. textbook. A CD-ROM comes with the book, including a trial version of the Risk Simulator and Real Options Super Lattice Solver software and associated Excel models.

JOHNATHAN MUN

San Francisco, California
JohnathanMun@cs.com
May 2006

Acknowledgments

The author is greatly indebted to Robert Fourt, Professor Morton Glantz, Dr. Charles Hardy, Steve Hoye, Professor Bill Rodney, Larry Pixley, Dr. Tom Housel, Lt. Commander Cesar Rios, Ken Cobleigh, Pat Haggerty, Larry Blair, Andy Roff, and Tony Jurado for their business case contributions. In addition, a special word of thanks goes to Bill Falloon, senior editor at John Wiley & Sons, for his support and encouragement.

J. M.

About the Author

Dr. Johnathan C. Mun is the founder and CEO of Real Options Valuation, Inc., a consulting, training, and software development firm specializing in real options, employee stock options, financial valuation, simulation, forecasting, optimization, and risk analysis located in northern California. He is the creator of the Real Options Super Lattice Solver software, Risk Simulator software, and Employee Stock Options Valuation software at the firm, as well as the risk-analysis training DVD, and he also holds public seminars on risk analysis and Certified Risk Analyst (CRA) programs. The Real Options Super Lattice Solver software showcased in this book supersedes the previous Real Options Analysis Toolkit software, which he also developed. He has also authored numerous other books including *Real Options Analysis: Tools and Techniques,* first and second editions (Wiley 2003 and 2005), *Real Options Analysis Course: Business Cases* (Wiley 2003), *Applied Risk Analysis: Moving Beyond Uncertainty* (Wiley 2003), *Valuing Employee Stock Options* (Wiley 2004), and others. His books and software are being used around the world at top universities (including the Bern Institute in Germany, Chung-Ang University in South Korea, Georgetown University, ITESM in Mexico, Massachusetts Institute of Technology, New York University, Stockholm University in Sweden, University of the Andes in Chile, University of Chile, University of Pennsylvania Wharton School, University of York in the United Kingdom, and Edinburgh University in Scotland, among others).

Dr. Mun is also currently a finance and economics professor and has taught courses in financial management, investments, real options, economics, and statistics at the undergraduate and the graduate M.B.A. levels. He is teaching and has taught at universities all over the world, from the U.S. Naval Postgraduate School (Monterey, California) and University of Applied Sciences (Switzerland and Germany) as full professor, to Golden Gate University (California) and St. Mary's College (California), and has chaired many graduate research thesis committees. He also teaches risk analysis, real options analysis, and risk for managers public courses where participants can obtain the Certified Risk Analyst (CRA) designation on completion of the week-long program. He was formerly the vice president of analytics at Decisioneering, Inc., where he headed up the development of real options and financial analytics software products, analytical consulting, training, and technical support, and where he was the creator of the Real Options Analysis Toolkit software, the older predecessor of the Real Options Super Lattice Software discussed in this book. Prior to joining Decisioneering, he was a consulting manager and financial economist in the Valuation Services and Global Financial Services practice of KPMG Consulting and a manager with the Economic Consulting Services practice at KPMG LLP. He has extensive experience in econometric modeling, financial analysis, real options, economic analysis, and statistics. During his tenure at Real Options Valuation, Inc., Decisioneering, and at KPMG Consulting, he had consulted on many real options, risk analysis, financial forecasting, project management, and financial valuation projects for multinational firms (current and former clients include 3M, Airbus, Boeing, BP, Chevron Texaco, Financial Accounting Standards Board, Fujitsu, GE, Microsoft, Motorola, U.S. Department of Defense, U.S. Navy, Veritas, and many others). His experience prior to joining KPMG included being department head of financial planning and analysis at Viking Inc. of FedEx, performing financial forecasting, economic analysis, and market research. Prior to that, he had also performed some financial planning and freelance financial consulting work.

Dr. Mun received a Ph.D. in finance and economics from Lehigh University, where his research and academic interests were in the areas of investment finance, econometric modeling, financial options, corporate finance, and microeconomic theory. He also has an M.B.A. in business administration, an M.S. in management science, and a B.S. in biology and physics. He is Certified in Financial Risk Management (FRM), Certified in Financial Consulting (CFC), and Certified in Risk Analysis (CRA). He is a member of the American Mensa, Phi Beta Kappa Honor Society, and Golden Key Honor Society as well as several other professional organizations, including the Eastern and Southern Finance Associations, American Economic Association, and Global Association of Risk Professionals. Finally, he has written many academic articles published in the *Journal of the Advances in Quantitative Accounting and Finance*, the *Global Finance Journal*, the *International Financial Review*, the *Journal of Financial Analysis*, the *Journal of Applied Financial Economics*, the *Journal of International Financial Markets, Institutions and Money*, the *Financial Engineering News*, and the *Journal of the Society of Petroleum Engineers*.

Contents

Introduction

This book is divided into nine parts starting from a discussion of what risk is and how it is quantified, to how risk can be predicted, diversified, taken advantage of, hedged, and, finally, managed. The first part deals with *risk identification* where the different aspects of business risks are identified, including a brief historical view of how risk was evaluated in the past. The second part deals with *risk evaluation* explaining why disastrous ramifications may result if risk is not considered in business decisions. Part Three pertains to *risk quantification* and details how risk can be captured quantitatively through step-by-step applications of Monte Carlo simulation. Part Four deals with *industry applications* and examples of how risk analysis is applied in practical day-to-day issues in the oil and gas, pharmaceutical, financial planning, hospital risk management, and executive compensation problems. Part Five pertains to *risk prediction* where the uncertain and risky future is predicted using analytical time-series methods. Part Six deals with how *risk diversification* works when multiple projects exist in a portfolio. Part Seven's *risk mitigation* discussion deals with how a firm or management can take advantage of risk and uncertainty by implementing and maintaining flexibility in projects. Part Eight provides a second installment of *business cases* where risk analysis is applied in the banking, real estate, military strategy, automotive parts aftermarket, and global earth observation systems. Part Nine provides a capstone discussion of applying *risk management* in companies, including how to obtain senior management's buy-in and implementing a change of perspective in corporate culture as it applies to risk analysis. This book is an update of *Applied Risk Analysis* (Wiley, 2004) to include coverage of the author's own Risk Simulator software and Real Options Super Lattice Solver software. Following is a synopsis of the material covered in each chapter of the book.

PART ONE—RISK IDENTIFICATION

Chapter 1—Moving Beyond Uncertainty

To the people who lived centuries ago, risk was simply the inevitability of chance occurrence beyond the realm of human control. We have been struggling with risk our entire existence, but, through trial and error and through the evolution of human knowledge and thought, have devised ways to describe and quantify risk. Risk assessment should be an important part of the decision-making process; otherwise bad decisions may be made. Chapter 1 explores the different facets of risk within the realms of applied business risk analysis, providing an intuitive feel of what risk is.

PART TWO—RISK EVALUATION

Chapter 2—From Risk to Riches

The concepts of risk and return are detailed in Chapter 2, illustrating their relationships in the financial world, where a higher-risk project necessitates a higher expected return. How are uncertainties estimated and risk calculated? How do you convert a measure of uncertainty into a measure of risk? These are the topics covered in this chapter, starting from the basics of statistics to applying them in risk analysis, and including a discussion of the different measures of risk.

Chapter 3—A Guide to Model-Building Etiquette

Chapter 3 addresses some of the more common errors and pitfalls analysts make when creating a new model by explaining some of the proper modeling etiquettes. The issues discussed range from file naming conventions and proper model aesthetics to complex data validation and Visual Basic for Applications (VBA) scripting. An appendix is provided on some VBA modeling basics and techniques of macros and forms creation.

PART THREE—RISK QUANTIFICATION

Chapter 4—On the Shores of Monaco

Monte Carlo simulation in its simplest form is just a random number generator useful for forecasting, estimation, and risk analysis. A simulation calculates numerous scenarios of a model by repeatedly picking values from the probability distribution for the uncertain variables and using those values for the event—events such as totals, net profit, or gross expenses. Simplistically, think of the Monte Carlo simulation approach as repeatedly picking golf balls out of a large basket. Chapter 4 illustrates why simulation is important through the flaw of averages example. Excel is used to perform rudimentary simulations, and simulation is shown as a logical next step extension to traditional approaches used in risk analysis.

Chapter 5—Test Driving Risk Simulator

Chapter 5 guides the user through applying the world's premier risk analysis and simulation software: *Risk Simulator*. With a few simple mouse clicks, the reader will be on his or her way to running sophisticated Monte Carlo simulation analysis to capture both uncertainty and risks using the enclosed CD-ROM's Risk Simulator trial software. In addition, the interpretation of said analysis is also very important. The best analysis in the world is only as good as the analyst's ability to understand, utilize, present, report, and convince management or clients of the results.

Chapter 6—Pandora's Toolbox

Powerful simulation-related tools such as bootstrapping, distributional fitting, hypothesis test, correlated simulation, multidimensional simulation, tornado charts, and sensitivity charts are discussed in detail in Chapter 6, complete with step-by-step illustrations. These tools are extremely valuable to analysts working in the realm of risk analysis. The applicability of each tool is discussed in detail. For example, the use of nonparametric bootstrapping simulation as opposed to parametric Monte Carlo simulation approaches is discussed. An appendix to this chapter deals with the technical specifics of goodness-of-fit tests.

PART FOUR—INDUSTRY APPLICATIONS

Chapter 7—Extended Business Cases I: Pharmaceutical and Biotech Negotiations, Oil and Gas Exploration, Financial Planning with Simulation, Hospital Risk Management, and Risk-Based Executive Compensation Valuation

Chapter 7 contains the first installment of actual business cases from industry applying risk analytics. Business cases were contributed by a variety of industry experts on applying risk analysis in the areas of oil and gas exploration, pharmaceutical biotech deal making, financial planning, hospital risk management, and executive compensation valuation.

PART FIVE—RISK PREDICTION

Chapter 8—Tomorrow's Forecast Today

Chapter 8 focuses on applying Risk Simulator to run time-series forecasting methods, multivariate regressions, nonlinear extrapolation, stochastic process forecasts, and Box-Jenkins ARIMA. In addition, the issues of seasonality and trend are discussed, together with the eight time-series decomposition models most commonly used by analysts to forecast future events given historical data. The software applications of each method are discussed in detail, complete with their associated measures of forecast errors and potential pitfalls.

Chapter 9—Using the Past to Predict the Future

The main thrust of Chapter 9 is time-series and regression analysis made easy. Starting with some basic time-series models, including exponential smoothing and moving averages, and moving on to more complex models, such as the Holt–Winters' additive and multiplicative models, the reader will manage to navigate through the maze of time-series analysis. The basics of regression analysis are also discussed, complete with pragmatic discussions of statistical validity tests as well as the pitfalls of regression analysis, including how to identify and fix heteroskedasticity, multicollinearity, and autocorrelation. The five appendixes that accompany this chapter deal with the technical specifics of interval estimations in regression analysis, ordinary least squares, and some pitfalls in running regressions, including detecting and fixing heteroskedasticity, multicollinearity, and autocorrelation.

PART SIX—RISK DIVERSIFICATION

Chapter 10—The Search for the Optimal Decision

In most business or analytical models, there are variables over which you have control, such as how much to charge for a product or how much to invest in a project. These controlled variables are called *decision variables*. Finding the optimal values for decision variables can make the difference between reaching an important goal and missing that goal. Chapter 10 details the optimization process at a high level, with illustrations on solving deterministic optimization problems manually, using graphs, and applying Excel's Solver add-in. (Chapter 11 illustrates the solution to optimization problems under uncertainty, mirroring more closely real-life business conditions.)

Chapter 11—Optimization Under Uncertainty

Chapter 11 illustrates two optimization models with step-by-step details. The first model is a discrete portfolio optimization of projects under uncertainty. Given a set of potential projects, the model evaluates all possible discrete combinations of projects on a "go" or "no-go" basis such that a budget constraint is satisfied, while simultaneously providing the best level of returns subject to uncertainty. The best projects will then be chosen based on these criteria. The second model evaluates a financial portfolio's continuous allocation of different asset classes with different levels of risks and returns. The objective of this model is to find the optimal allocation of assets subject to a 100 percent allocation constraint that still maximizes the Sharpe ratio, or the portfolio's return-to-risk ratio. This ratio will maximize the portfolio's return subject to the minimum risks possible while accounting for the cross-correlation diversification effects of the asset classes in a portfolio.

PART SEVEN—RISK MITIGATION

Chapter 12—What Is So Real about Real Options, and Why Are They Optional?

Chapter 12 describes what real option analysis is, who has used the approach, how companies are using it, and what some of the characteristics of real options are. The chapter describes real options in a nutshell, providing the reader with a solid introduction to its concepts without the need for its theoretical underpinnings. Real options are applicable if the following requirements are met: traditional financial analysis can be performed and models can be built; uncertainty exists; the same uncertainty drives value; management or the project has strategic options or flexibility to either take advantage of these uncertainties or to hedge them; and management must be credible to execute the relevant strategic options when they become optimal to do so.

Chapter 13—The Black Box Made Transparent: Real Options Super Lattice Solver Software

Chapter 13 introduces the readers to the world's first true real options software applicable across all industries. The chapter illustrates how a user can get started with the software in a few short moments after it has been installed. The reader is provided with hands-on experience with the Real Options Super Lattice Solver to obtain immediate results—a true test when the rubber meets the road.

PART EIGHT—MORE INDUSTRY APPLICATIONS

Chapter 14—Extended Business Cases II: Real Estate, Banking, Military Strategy, Automotive Aftermarkets, Global Earth Observing Systems, and Valuing Employee Stock Options (FAS 123R)

Chapter 14 contains the second installment of actual business cases from industry applying risk analytics. Business cases were contributed by a variety of industry experts applying simulation, optimization, and real options analysis in the areas of real estate, banking, military strategy, automotive parts aftermarket, global earth observing systems, and employee stock options.

PART NINE—RISK MANAGEMENT

Chapter 15—The Warning Signs

The risk analysis software applications illustrated in this book are extremely powerful tools and could prove detrimental in the hands of untrained and unlearned novices. Management, the end user of the results from said tools, must be able to discern if quality analysis has been performed. Chapter 15 delves into the thirty-some problematic issues most commonly encountered by analysts applying risk analysis techniques, and how management can spot these mistakes. While it might be the job of the analyst to create the models and use the fancy analytics, it is senior management's job to challenge the assumptions and results obtained from the analysis. Model errors, assumption and input errors, analytical errors, user errors, and interpretation errors are some of the issues discussed in this chapter. Some of the issues and concerns raised for management's consideration in performing due diligence include challenging distributional assumptions, critical success factors, impact drivers, truncation, forecast validity, endpoints, extreme values, structural breaks, values at risk, a priori expectations, back-casting, statistical validity, specification errors, out of range forecasts, heteroskedasticity, multicollinearity, omitted variables, spurious relationships, causality and correlation, autoregressive processes, seasonality, random walks, and stochastic processes.

Chapter 16—Changing a Corporate Culture

Advanced analytics is hard to explain to management. So, how do you get risk analysis accepted as the norm into a corporation, especially if your industry is highly conservative? It is a guarantee in companies like these that an analyst showing senior management a series of fancy and mathematically sophisticated models will be thrown out of the office together with his or her results, and have the door slammed shut. Change management is the topic of discussion in Chapter 16. Explaining the results and convincing management appropriately go hand in hand with the characteristics of the analytical tools, which, if they satisfy certain change management requisites, can make acceptance easier. The approach that guarantees acceptance has to be three pronged: Top, middle, and junior levels must all get in on the action. Change management specialists underscore that change comes more easily if the methodologies to be accepted are applicable to the problems at hand, are accurate and consistent, provide value-added propositions, are easy to explain, have comparative advantage over traditional approaches, are compatible with the old, have modeling flexibility, are backed by executive sponsorship, and are influenced and championed by external parties including competitors, customers, counterparties, and vendors.

ADDITIONAL MATERIAL

The book concludes with the ten mathematical tables used in the analyses throughout the book and the answers to the questions and exercises at the end of each chapter. The CD-ROM included with the book holds 30-day trial versions of Risk Simulator and Real Options Super Lattice Solver software, as well as sample models and getting started videos to help the reader get a jump start on modeling risk.

PART One

Risk Identification

CHAPTER 1

Moving Beyond Uncertainty

A BRIEF HISTORY OF RISK: WHAT EXACTLY IS RISK?

Since the beginning of recorded history, games of chance have been a popular pastime. Even in Biblical accounts, Roman soldiers cast lots for Christ's robes. In earlier times, chance was something that occurred in nature, and humans were simply subjected to it as a ship is to the capricious tosses of the waves in an ocean. Even up to the time of the Renaissance, the future was thought to be simply a chance occurrence of completely random events and beyond the control of humans. However, with the advent of games of chance, human greed has propelled the study of risk and chance to evermore closely mirror real-life events. Although these games initially were played with great enthusiasm, no one actually sat down and figured out the odds. Of course, the individual who understood and mastered the concept of chance was bound to be in a better position to profit from such games of chance. It was not until the mid-1600s that the concept of chance was properly studied, and the first such serious endeavor can be credited to Blaise Pascal, one of the fathers of modern choice, chance, and probability.[1] Fortunately for us, after many centuries of mathematical and statistical innovations from pioneers such as Pascal, Bernoulli, Bayes, Gauss, LaPlace, and Fermat, our modern world of uncertainty can be explained with much more elegance through methodological applications of risk and uncertainty.

To the people who lived centuries ago, risk was simply the inevitability of chance occurrence beyond the realm of human control. Nonetheless, many phony soothsayers profited from their ability to convincingly profess their clairvoyance by simply stating the obvious or reading the victims' body language and telling them what they wanted to hear. We modern-day humans, ignoring for the moment the occasional seers among us, with our fancy technological achievements, are still susceptible to risk and uncertainty. We may be able to predict the orbital paths of planets in our solar system with astounding accuracy or the escape velocity required to shoot a man from the Earth to the Moon, but when it comes to predicting a firm's revenues the following year, we are at a loss. Humans have been struggling with risk our entire existence but, through trial and error, and through the evolution of human knowledge and thought, have devised ways to describe, quantify, hedge, and take advantage of risk.

Clearly the entire realm of risk analysis is great and would most probably be intractable within the few chapters of a book. Therefore, this book is concerned with only a small niche of risk, namely *applied business risk modeling and analysis*. Even in the areas of applied business risk analysis, the diversity is great. For instance, business risk can be roughly divided into the areas of operational risk management and financial risk management. In financial risk, one can look at market risk, private risk, credit risk, default risk, maturity risk, liquidity risk, inflationary risk, interest rate risk, country risk, and so forth. This book focuses on the application of risk analysis in the sense of how to adequately apply the tools to identify, understand, quantify, and diversify risk such that it can be hedged and managed more effectively. These tools are generic enough that they can be applied across a whole spectrum of business conditions, industries, and needs.

Finally, understanding this text in its entirety together with *Real Options Analysis*, Second Edition (Wiley, 2005) and the associated Risk Simulator and Real Options SLS software are required prerequisites for the Certified Risk Analyst or CRA certification (see www.realoptionsvaluation.com for more details).

UNCERTAINTY VERSUS RISK

Risk and uncertainty are very different-looking animals, but they are of the same species; however, the lines of demarcation are often blurred. A distinction is critical at this juncture before proceeding and worthy of segue. Suppose I am senseless enough to take a skydiving trip with a good friend and we board a plane headed for the Palm Springs desert. While airborne at 10,000 feet and watching our lives flash before our eyes, we realize that in our haste we forgot to pack our parachutes on board. However, there is an old, dusty, and dilapidated emergency parachute on the plane. At that point, both my friend and I have the same level of uncertainty—the uncertainty of whether the old parachute will open, and if it does not, whether we will fall to our deaths. However, being the risk-adverse, nice guy I am, I decide to let my buddy take the plunge. Clearly, he is the one taking the plunge and the same person taking the risk. I bear no risk at this time while my friend bears all the risk.[2] However, we both have the same level of uncertainty as to whether the parachute will actually fail. In fact, we both have the same level of uncertainty as to the outcome of the day's trading on the New York Stock Exchange—which has absolutely no impact on whether we live or die that day. Only when he jumps and the parachute opens will the uncertainty become resolved through the passage of time, events, and action. However, even when the uncertainty is resolved with the opening of the parachute, the risk still exists as to whether he will land safely on the ground below.

Therefore, risk is something one bears and is the outcome of uncertainty. Just because there is uncertainty, there could very well be no risk. If the only thing that bothers a U.S.-based firm's CEO is the fluctuation in the foreign exchange market of the Zambian kwacha, then I might suggest shorting some kwachas and shifting his portfolio to U.S.-based debt. This uncertainty, if it does not affect the firm's bottom line in any way, is only uncertainty and not risk. This book is concerned with risk by performing uncertainty analysis—the same uncertainty that brings about risk by its mere existence as it impacts the value of a particular project. It is further assumed that the end user of this uncertainty analysis uses the results appropriately, whether the analysis is for identifying, adjusting, or selecting projects with respect to their risks, and so forth. Otherwise, running millions of fancy simulation trials and letting the results "marinate" will be useless. By running simulations on the foreign exchange market of the kwacha, an analyst sitting in a cubicle somewhere in downtown Denver will in no way reduce the risk of the kwacha in the market or the firm's exposure to the same. Only by using the results from an uncertainty simulation analysis and finding ways to hedge or mitigate the quantified fluctuation and downside risks of the firm's foreign exchange exposure through the derivatives market could the analyst be construed as having performed risk analysis and risk management.

To further illustrate the differences between risk and uncertainty, suppose we are attempting to forecast the stock price of Microsoft (MSFT). Suppose MSFT is currently priced at $25 per share, and historical prices place the stock at 21.89% volatility. Now suppose that for the next 5 years, MSFT does not engage in any risky ventures and stays exactly the way it is, and further suppose that the entire economic and financial world remains constant. This means that *risk* is fixed and unchanging; that is, volatility is unchanging for the next 5 years. However, the price uncertainty still increases over time; that is, the width of the forecast intervals will still increase over time. For instance, Year 0's forecast is known and is $25. However, as we progress one day, MSFT will most probably vary between $24 and $26. One year later, the uncertainty bounds may be between $20 and $30. Five years into the future, the boundaries might be between $10 and $50. So, in this example, *uncertainties increase* while *risks remain the same*. Therefore, risk is not equal to uncertainty. This idea is, of course, applicable to any forecasting approach whereby it becomes more and more difficult to forecast the future albeit the same risk. Now, if risk changes over time, the bounds of uncertainty get more complicated (e.g., uncertainty bounds of sinusoidal waves with discrete event jumps).

In other instances, risk and uncertainty are used interchangeably. For instance, suppose you play a coin-toss game—bet $0.50 and if heads come up you win $1, but you lose everything if tails appear. The risk here is you lose everything because the risk is that tails may appear. The uncertainty here is that tails may appear. Given that tails appear, you lose everything; hence, uncertainty brings with it risk. Uncertainty is the possibility of an event occurring and risk is the ramification of such an event occurring. People tend to use these two terms interchangeably.

In discussing uncertainty, there are three levels of uncertainties in the world: the *known*, the *unknown*, and the *unknowable*. The known is, of course, what we know will occur and are certain of its occurrence (contractual obligations or a guaranteed event); the unknown is what we do not know and can be simulated. These events will become known through the passage of time, events, and action (the uncertainty of whether a new drug or technology can be developed successfully will become known after spending years and millions on research programs—it will either work or not, and we will know this in the future), and these events carry with them risks, but these risks will be reduced or eliminated over time. However, unknowable events carry both uncertainty and risk that the totality of the risk and uncertainty may not change through the passage of time, events, or actions. These are events such as when the next tsunami or earthquake will hit, or when another act of terrorism will occur around the world. When an event occurs, uncertainty becomes resolved, but risk still remains (another one may or may not hit tomorrow). In traditional analysis, we care about the known factors. In risk analysis, we care about the unknown and unknowable factors. The unknowable factors are easy to hedge—get the appropriate insurance! That is, do not do business in a war-torn country, get away from politically unstable economies, buy hazard and business interruption insurance, and so forth. It is for the unknown factors that risk analysis will provide the most significant amount of value.

WHY IS RISK IMPORTANT IN MAKING DECISIONS?

Risk should be an important part of the decision-making process; otherwise bad decisions may be made without an assessment of risk. For instance, suppose projects are chosen based simply on an evaluation of returns; clearly the highest-return project will be chosen over lower-return projects. In financial theory, projects with higher returns will in most cases bear higher risks.[3] Therefore, instead of relying purely on bottom-line profits, a project should be evaluated based on its returns as well as its risks. Figures 1.1 and 1.2 illustrate the errors in judgment when risks are ignored.

> The concepts of risk and uncertainty are related but different. Uncertainty involves variables that are unknown and changing, but its uncertainty will become known and resolved through the passage of time, events, and action. Risk is something one bears and is the outcome of uncertainty. Sometimes, risk may remain constant while uncertainty increases over time.

Figure 1.1 lists three *mutually exclusive* projects with their respective-costs to implement, expected net returns (net of the costs to implement), and risk levels (all in present values).[4] Clearly, for the budget-constrained manager, the cheaper the project the better, resulting in the selection of Project X.[5] The returns-driven manager will choose Project Y with the highest returns, assuming that budget is not an issue. Project Z will be chosen by the risk-averse manager as it provides the least amount of risk while providing a positive net return. The upshot is that with three different projects and three different managers, three different decisions will be made. Which manager is correct and why?

Figure 1.2 shows that Project Z should be chosen. For illustration purposes, suppose all three projects are independent and mutually exclusive,[6] and that an unlimited number of projects from each category can be chosen but the budget is constrained at $1,000. Therefore, with this $1,000 budget, 20 project Xs can be chosen, yielding $1,000 in net returns and $500 risks, and so forth. It is clear from Figure 1.2 that project Z is the best project as for the same level of net returns ($1,000), the least amount of risk is undertaken ($100). Another way of viewing this selection is that for each $1 of returns obtained, only $0.1 amount of risk is involved on average, or that for each $1 of risk, $10 in returns are obtained on average. This example illustrates the concept of *bang for the buck* or getting the best value with the

Name of Project	Cost	Returns	Risk
Project X	$50	$50	$25
Project Y	$250	$200	$200
Project Z	$100	$100	$10

Project X for the cost- and budget-constrained manager
Project Y for the returns-driven and nonresource-constrained manager
Project Z for the risk-averse manager
Project Z for the smart manager

FIGURE 1.1 Why is risk important?

Looking at bang for the buck, X (2), Y (1), Z (10), Project Z should be chosen — with a $1,000 budget, the following can be obtained:

Project X: 20 Project Xs returning $1,000, with $500 risk
Project Y: 4 Project Xs returning $800, with $800 risk
Project Z: 10 Project Xs returning $1,000, with $100 risk

Project X: For each $1 return, $0.5 risk is taken
Project Y: For each $1 return, $1.0 risk is taken
Project Z: For each $1 return, $0.1 risk is taken

Project X: For each $1 of risk taken, $2 return is obtained
Project Y: For each $1 of risk taken, $1 return is obtained
Project Z: For each $1 of risk taken, $10 return is obtained

Conclusion: Risk is important. Ignoring risks results in making the wrong decision.

FIGURE 1.2 Adding an element of risk.

least amount of risk. An even more blatant example is if there are several different projects with identical single-point average net returns of $10 million each. Without risk analysis, a manager should in theory be indifferent in choosing any of the projects.[7] However, with risk analysis, a better decision can be made. For instance, suppose the first project has a 10 percent chance of exceeding $10 million, the second a 15 percent chance, and the third a 55 percent chance. The third project, therefore, is the best bet.

DEALING WITH RISK THE OLD-FASHIONED WAY

Businesses have been dealing with risk since the beginning of the history of commerce. In most cases, managers have looked at the risks of a particular project, acknowledged their existence, and moved on. Little quantification was performed in the past. In fact, most decision makers look only to single-point estimates of a project's profitability. Figure 1.3 shows an example of a single-point estimate. The estimated net revenue of $30 is simply that, a single point whose probability of occurrence is close to zero.[8] Even in the simple model shown in Figure 1.3, the effects of interdependencies are ignored, and in traditional modeling jargon, we have the problem of *garbage in, garbage out* (GIGO). As an example of interdependencies, the units sold are probably negatively correlated to the price of the product,[9] and positively correlated to the average variable cost;[10] ignoring these effects in a single-point estimate will yield grossly incorrect results. For instance, if the unit sales variable becomes 11 instead of 10, the resulting revenue may not

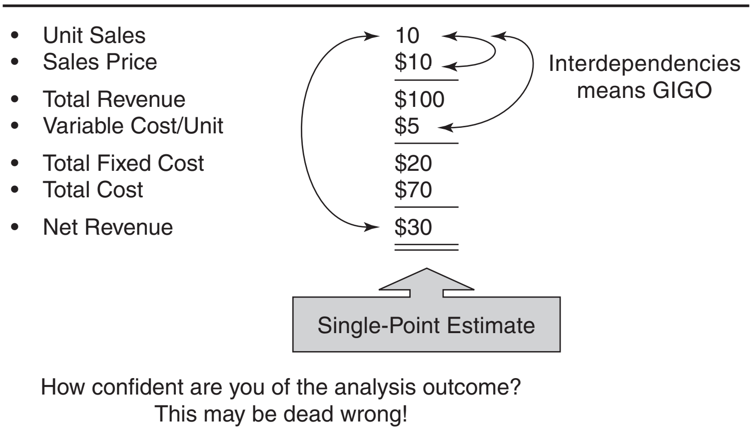


FIGURE 1.3 Single-point estimate.

simply be $35. The net revenue may actually decrease due to an increase in variable cost per unit while the sale price may actually be slightly lower to accommodate this increase in unit sales. Ignoring these interdependencies will reduce the accuracy of the model.

> A rational manager would choose projects based not only on returns but also on risks. The best projects tend to be those with the best bang for the buck, or the best returns subject to some specified risks.

One approach used to deal with risk and uncertainty is the application of scenario analysis, as seen in Figure 1.4. Suppose the worst-case, nominal-case, and best-case scenarios are applied to the unit sales; the resulting three scenarios' net revenues are obtained. As earlier, the problems of interdependencies are not addressed. The net revenues obtained are simply too variable, ranging from $5 to $55. Not much can be determined from this analysis.

A related approach is to perform *what-if* or *sensitivity* analysis as seen in Figure 1.5. Each variable is perturbed and varied a prespecified amount and the resulting change in net revenues is captured. This approach is great for understanding which variables drive or impact the bottom line the most. A related approach is the use of tornado and sensitivity charts as detailed in Chapter 6, Pandora's Toolbox, which looks at a series of simulation tools. These approaches were usually the extent to which risk and uncertainty

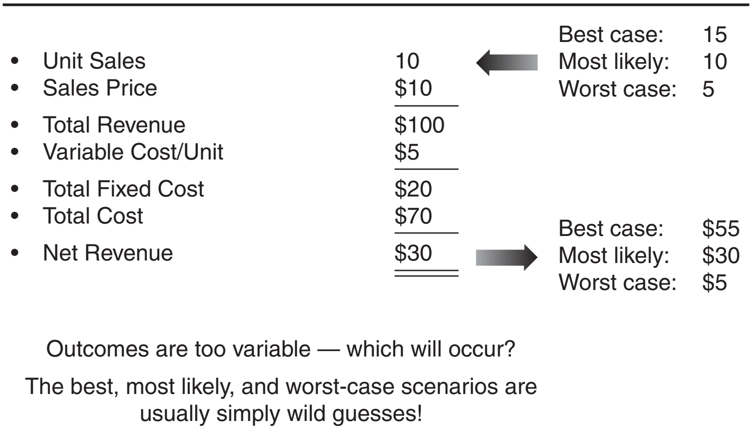


FIGURE 1.4 Scenario analysis.

analysis were traditionally performed. Clearly, a better and more robust approach is required.

This is the point where simulation comes in. Figure 1.6 shows how simulation can be viewed as simply an extension of the traditional approaches of sensitivity and scenario testing. The critical success drivers or the variables that affect the bottom-line net-revenue variable the most, which at the same time are uncertain, are simulated. In simulation, the interdependencies are accounted for by using correlations. The uncertain variables are then simulated thousands of times to emulate all potential permutations and combi-

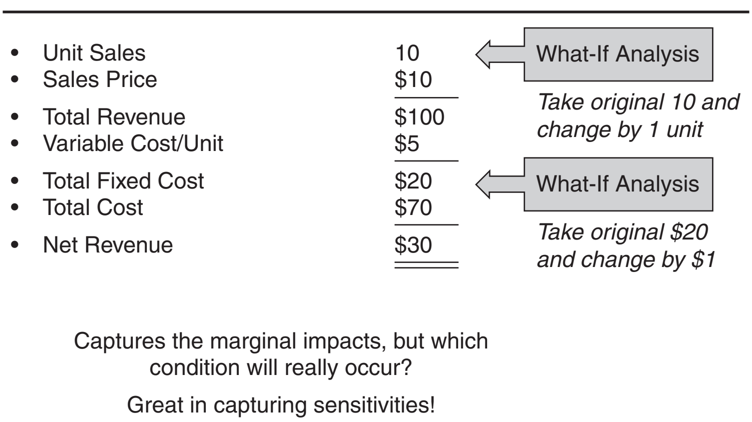


FIGURE 1.5 What-if sensitivity analysis.

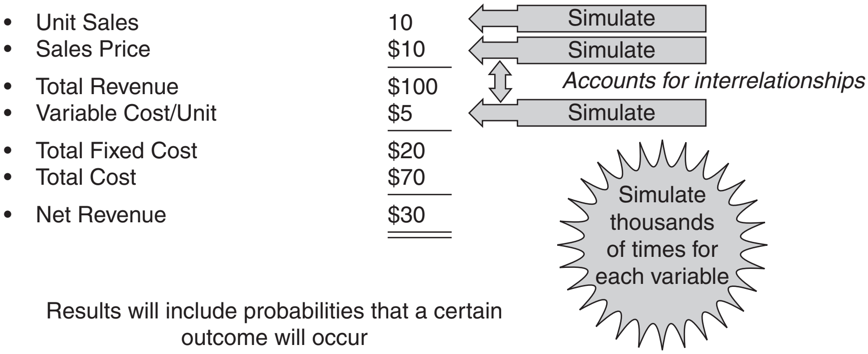


FIGURE 1.6 Simulation approach.

nations of outcomes. The resulting net revenues from these simulated potential outcomes are tabulated and analyzed. In essence, in its most basic form, simulation is simply an enhanced version of traditional approaches such as sensitivity and scenario analysis but automatically performed for thousands of times while accounting for all the dynamic interactions between the simulated variables. The resulting net revenues from simulation, as seen in Figure 1.7, show that there is a 90 percent probability that the net

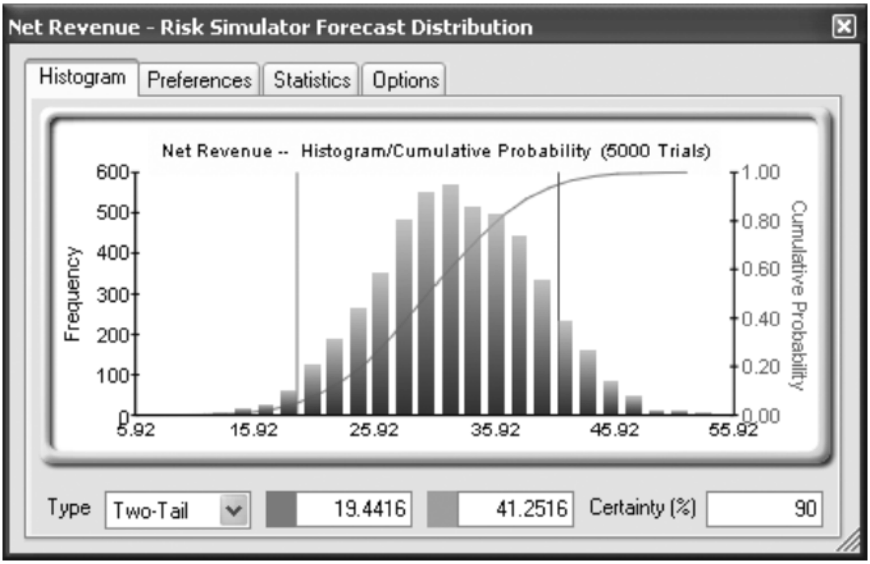


FIGURE 1.7 Simulation results.

revenues will fall between \$19.44 and \$41.25, with a 5 percent worst-case scenario of net revenues falling below \$19.44. Rather than having only three scenarios, simulation created 5,000 scenarios, or trials, where multiple variables are simulated and changing simultaneously (unit sales, sale price, and variable cost per unit), while their respective relationships or correlations are maintained.

THE LOOK AND FEEL OF RISK AND UNCERTAINTY

In most financial risk analyses, the first step is to create a series of free cash flows (FCF), which can take the shape of an income statement or discounted cash-flow (DCF) model. The resulting deterministic free cash flows are depicted on a time line, akin to that shown in Figure 1.8. These cash-flow figures are in most cases forecasts of the unknown future. In this simple example, the cash flows are assumed to follow a straight-line growth curve (of course, other shaped curves also can be constructed). Similar forecasts

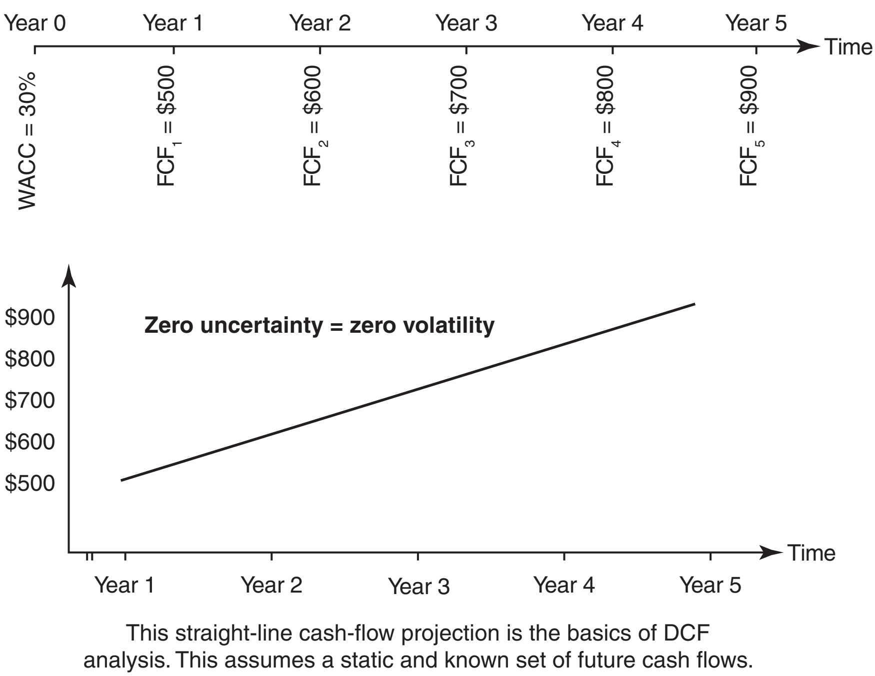
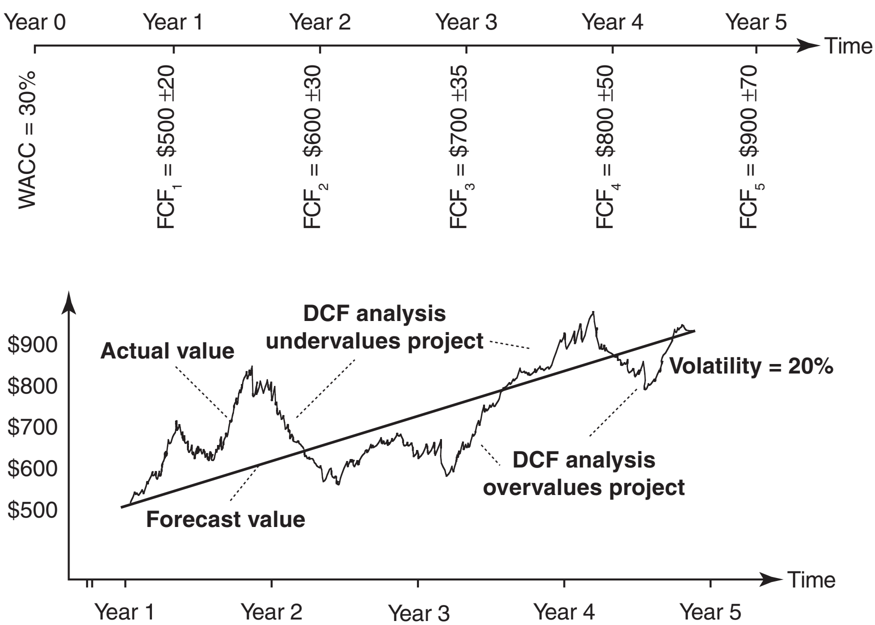


FIGURE 1.8 The intuition of risk—deterministic analysis.

can be constructed using historical data and fitting these data to a time-series model or a regression analysis.[11] Whatever the method of obtaining said forecasts or the shape of the growth curve, these are point estimates of the unknown future. Performing a financial analysis on these static cash flows provides an accurate value of the project if and only if all the future cash flows are known with certainty—that is, no uncertainty exists.

However, in reality, business conditions are hard to forecast. Uncertainty exists, and the actual levels of future cash flows may look more like those in Figure 1.9; that is, at certain time periods, actual cash flows may be above, below, or at the forecast levels. For instance, at any time period, the actual cash flow may fall within a range of figures with a certain percent probability. As an example, the first year's cash flow may fall anywhere between \$480 and \$520. The actual values are shown to fluctuate around the forecast values at an average volatility of 20 percent.[12] (We use volatility here as a measure of uncertainty, i.e., the higher the volatility, the higher the level of uncertainty, where at zero uncertainty, the outcomes are 100 percent certain[13]). Certainly this example provides a much more accurate view of the



This graph shows that in reality, at different times, actual cash flows may be above, below, or at the forecast value line due to uncertainty and risk.

FIGURE 1.9 The intuition of risk—Monte Carlo simulation.

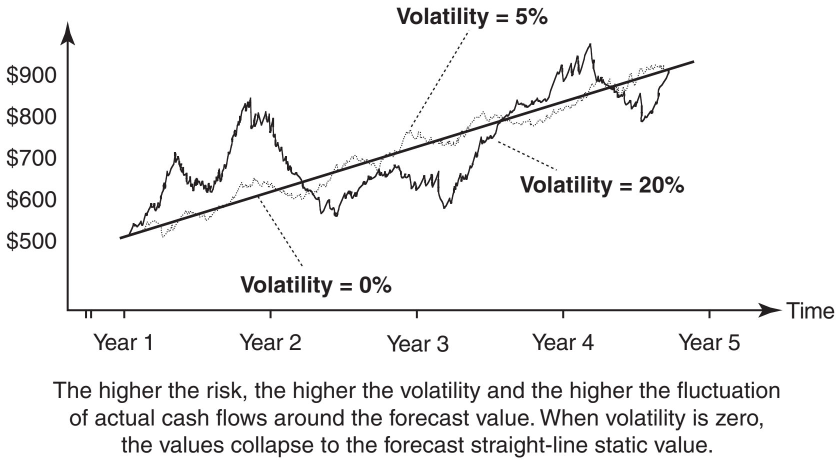


FIGURE 1.10 The intuition of risk—the face of risk.

true nature of business conditions, which are fairly difficult to predict with any amount of certainty.

Figure 1.10 shows two sample actual cash flows around the straight-line forecast value. The higher the uncertainty around the actual cash-flow levels, the higher the volatility. The darker line with 20 percent volatility fluctuates more wildly around the forecast values. These values can be quantified using Monte Carlo simulation fairly easily but cannot be properly accounted for using more simplistic traditional methods such as sensitivity or scenario analyses.

INTEGRATED RISK ANALYSIS FRAMEWORK

Before diving into the different risk analysis methods in the remaining chapters of the book, it is important to first understand the *integrated risk analysis framework* and how these different techniques are related in a risk analysis and risk management context. This framework comprises eight distinct phases of a successful and comprehensive risk analysis implementation, going from a qualitative management screening process to creating clear and concise reports for management. The process was developed by the author based on previous successful implementations of risk analysis, forecasting, real options, valuation, and optimization projects both in the consulting arena and in industry-specific problems. These phases can be performed either in isolation or together in sequence for a more robust integrated analysis.

Figure 1.11 shows the integrated risk analysis process up close. We can segregate the process into the following eight simple steps:

1. Qualitative management screening.
2. Time-series and regression forecasting.
3. Base case net present value analysis.
4. Monte Carlo simulation.
5. Real options problem framing.
6. Real options modeling and analysis.
7. Portfolio and resource optimization.
8. Reporting and update analysis.

1. Qualitative Management Screening

Qualitative management screening is the first step in any integrated risk analysis process. Management has to decide which projects, assets, initiatives, or strategies are viable for further analysis, in accordance with the firm's mission, vision, goal, or overall business strategy. The firm's mission, vision, goal, or overall business strategy may include market penetration strategies, competitive advantage, technical, acquisition, growth, synergistic, or globalization issues. That is, the initial list of projects should be qualified in terms of meeting management's agenda. Often at this point the most valuable insight is created as management frames the complete problem to be resolved and the various risks to the firm are identified and flushed out.

2. Time-Series and Regression Forecasting

The future is then forecasted using time-series analysis or multivariate regression analysis if historical or comparable data exist. Otherwise, other qualitative forecasting methods may be used (subjective guesses, growth rate assumptions, expert opinions, Delphi method, and so forth). In a financial context, this is the step where future revenues, sale price, quantity sold, volume, production, and other key revenue and cost drivers are forecasted. See Chapters 8 and 9 for details on forecasting and using the author's Risk Simulator software to run time-series, extrapolation, stochastic process, ARIMA, and regression forecasts.

3. Base Case Net Present Value Analysis

For each project that passes the initial qualitative screens, a discounted cash flow model is created. This model serves as the base case analysis where a net present value (NPV) is calculated for each project, using the forecasted values from the previous step. This step also applies if only a single project is under evaluation. This net present value is calculated using the traditional

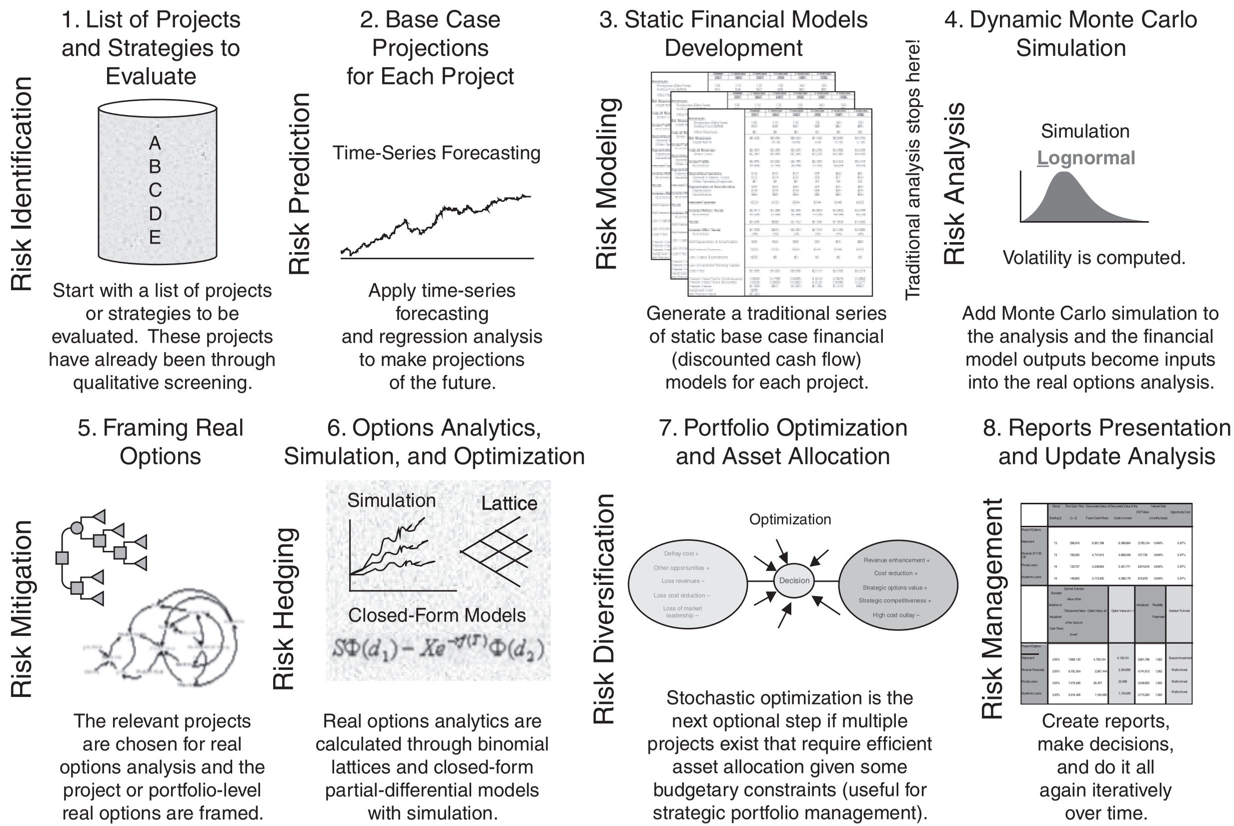


FIGURE 1.11 Integrated risk analysis process.

approach of using the forecast revenues and costs, and discounting the net of these revenues and costs at an appropriate risk-adjusted rate. The return on investment and other metrics are generated here.

4. Monte Carlo Simulation

Because the static discounted cash flow produces only a single-point estimate result, there is oftentimes little confidence in its accuracy given that future events that affect forecast cash flows are highly uncertain. To better estimate the actual value of a particular project, Monte Carlo simulation should be employed next. See Chapters 4 and 5 for details on running Monte Carlo simulations using the author's Risk Simulator software.

Usually, a sensitivity analysis is first performed on the discounted cash flow model; that is, setting the net present value as the resulting variable, we can change each of its precedent variables and note the change in the resulting variable. Precedent variables include revenues, costs, tax rates, discount rates, capital expenditures, depreciation, and so forth, which ultimately flow through the model to affect the net present value figure. By tracing back all these precedent variables, we can change each one by a preset amount and see the effect on the resulting net present value. A graphical representation can then be created, which is often called a tornado chart (see Chapter 6 on using Risk Simulator's simulation analysis tools such as tornado charts, spider charts, and sensitivity charts), because of its shape, where the most sensitive precedent variables are listed first, in descending order of magnitude. Armed with this information, the analyst can then decide which key variables are highly uncertain in the future and which are deterministic. The uncertain key variables that drive the net present value and, hence, the decision are called critical success drivers. These critical success drivers are prime candidates for Monte Carlo simulation. Because some of these critical success drivers may be correlated—for example, operating costs may increase in proportion to quantity sold of a particular product, or prices may be inversely correlated to quantity sold—a correlated Monte Carlo simulation may be required. Typically, these correlations can be obtained through historical data. Running correlated simulations provides a much closer approximation to the variables' real-life behaviors.

5. Real Options Problem Framing

The question now is that after quantifying risks in the previous step, what next? The risk information obtained somehow needs to be converted into *actionable intelligence*. Just because risk has been quantified to be such and such using Monte Carlo simulation, so what, and what do we do about it? The answer is to use real options analysis to hedge these risks, to value these risks, and to position yourself to take advantage of the risks. The first step in real options is to generate a strategic map through the process of framing the problem. Based on the overall problem identification occurring during the initial qualitative management screening process, certain strategic optionalities would have become apparent for each particular project. The strategic optionalities may include, among other things, the option to expand, contract, abandon, switch, choose, and so forth. Based on the identification of strategic optionalities that exist for each project or at each stage of the project, the analyst can then choose from a list of options to analyze in more detail. Real options are added to the projects to hedge downside risks and to take advantage of upside swings.

6. Real Options Modeling and Analysis

Through the use of Monte Carlo simulation, the resulting stochastic discounted cash flow model will have a distribution of values. Thus, simulation models, analyzes, and quantifies the various risks and uncertainties of each project. The result is a distribution of the NPVs and the project's volatility. In real options, we assume that the underlying variable is the future profitability of the project, which is the future cash flow series. An implied volatility of the future free cash flow or underlying variable can be calculated through the results of a Monte Carlo simulation previously performed. Usually, the volatility is measured as the standard deviation of the logarithmic returns on the free cash flow stream. In addition, the present value of future cash flows for the base case discounted cash flow model is used as the initial underlying asset value in real options modeling. Using these inputs, real options analysis is performed to obtain the projects' strategic option values—see Chapters 12 and 13 for details on understanding the basics of real options and on using the Real Options Super Lattice Solver software.

7. Portfolio and Resource Optimization

Portfolio optimization is an optional step in the analysis. If the analysis is done on multiple projects, management should view the results as a portfolio of rolled-up projects because the projects are in most cases correlated with one another, and viewing them individually will not present the true picture. As firms do not only have single projects, portfolio optimization is crucial. Given that certain projects are related to others, there are opportunities for hedging and diversifying risks through a portfolio. Because firms have limited budgets, have time and resource constraints, while at the same time have requirements for certain overall levels of returns, risk tolerances, and so forth, portfolio optimization takes into account all these to create an optimal portfolio mix. The analysis will provide the optimal allocation of investments across multiple projects. See Chapters 10 and 11 for details on using Risk Simulator to perform portfolio optimization.

8. Reporting and Update Analysis

The analysis is not complete until reports can be generated. Not only are results presented, but the process should also be shown. Clear, concise, and precise explanations transform a difficult black-box set of analytics into transparent steps. Management will never accept results coming from black boxes if they do not understand where the assumptions or data originate and what types of mathematical or financial massaging takes place.

Risk analysis assumes that the future is uncertain and that management has the right to make midcourse corrections when these uncertainties become resolved or risks become known; the analysis is usually done ahead of time and, thus, ahead of such uncertainty and risks. Therefore, when these risks become known, the analysis should be revisited to incorporate the decisions made or revising any input assumptions. Sometimes, for long-horizon projects, several iterations of the real options analysis should be performed, where future iterations are updated with the latest data and assumptions.

Understanding the steps required to undertake an integrated risk analysis is important because it provides insight not only into the methodology itself, but also into how it evolves from traditional analyses, showing where the traditional approach ends and where the new analytics start.

QUESTIONS

1. Why is risk important in making decisions?
2. Describe the concept of bang for the buck.
3. Compare and contrast risk and uncertainty.

PART Two

Risk Evaluation

CHAPTER 2
From Risk to Riches

TAMING THE BEAST

Risky ventures are the norm in the daily business world. The mere mention of names such as George Soros, John Meriweather, Paul Reichmann, and Nicholas Leeson, or firms such as Long Term Capital Management, Metallgesellschaft, Barings Bank, Bankers Trust, Daiwa Bank, Sumimoto Corporation, Merrill Lynch, and Citibank brings a shrug of disbelief and fear. These names are some of the biggest in the world of business and finance. Their claim to fame is not simply being the best and brightest individuals or being the largest and most respected firms, but for bearing the stigma of being involved in highly risky ventures that turned sour almost overnight.[1]

George Soros was and still is one of the most respected names in high finance; he is known globally for his brilliance and exploits. Paul Reichmann was a reputable and brilliant real estate and property tycoon. Between the two of them, nothing was impossible, but when they ventured into investments in Mexican real estate, the wild fluctuations of the peso in the foreign exchange market was nothing short of a disaster. During late 1994 and early 1995, the peso hit an all-time low and their ventures went from bad to worse, but the one thing that they did not expect was that the situation would become a lot worse before it was all over and billions would be lost as a consequence.

Long Term Capital Management was headed by Meriweather, one of the rising stars in Wall Street, with a slew of superstars on its management team, including several Nobel laureates in finance and economics (Robert Merton and Myron Scholes). The firm was also backed by giant investment banks. A firm that seemed indestructible literally blew up with billions of dollars in the red, shaking the international investment community with repercussions throughout Wall Street as individual investors started to lose faith in large hedge funds and wealth-management firms, forcing the eventual massive Federal Reserve bailout.

Barings was one of the oldest banks in England. It was so respected that even Queen Elizabeth II herself held a private account with it. This multibillion dollar institution was brought down single-handedly by Nicholas Leeson, an employee halfway around the world. Leeson was a young and brilliant investment banker who headed up Barings' Singapore branch. His illegally doctored track record showed significant investment profits, which gave him more leeway and trust from the home office over time. He was able to cover his losses through fancy accounting and by taking significant amounts of risk. His speculations in the Japanese yen went south and he took Barings down with him, and the top echelon in London never knew what hit them.

Had any of the managers in the boardroom at their respective headquarters bothered to look at the risk profile of their investments, they would surely have made a very different decision much earlier on, preventing what became major embarrassments in the global investment community. If the projected returns are adjusted for risks, that is, finding what levels of risks are required to attain such seemingly extravagant returns, it would be sensible not to proceed.

Risks occur in everyday life that do not require investments in the multimillions. For instance, when would one purchase a house in a fluctuating housing market? When would it be more profitable to lock in a fixed-rate mortgage rather than keep a floating variable rate? What are the chances that there will be insufficient funds at retirement? What about the potential personal property losses when a hurricane hits? How much accident insurance is considered sufficient? How much is a lottery ticket actually worth?

Risk permeates all aspects of life and one can never avoid taking or facing risks. What we can do is understand risks better through a systematic assessment of their impacts and repercussions. This assessment framework must also be capable of measuring, monitoring, and managing risks; otherwise, simply noting that risks exist and moving on is not optimal. This book provides the tools and framework necessary to tackle risks head-on. Only with the added insights gained through a rigorous assessment of risk can one actively manage and monitor risk.

> Risks permeate every aspect of business, but we do not have to be passive participants. What we can do is develop a framework to better understand risks through a systematic assessment of their impacts and repercussions. This framework also must be capable of measuring, monitoring, and managing risks.

THE BASICS OF RISK

Risk can be defined simply as any uncertainty that affects a system in an unknown fashion whereby the ramifications are also unknown but bears with it great fluctuation in value and outcome. In every instance, for risk to be evident, the following generalities must exist:

- Uncertainties and risks have a time horizon.
- Uncertainties exist in the future and will evolve over time.
- Uncertainties become risks if they affect the outcomes and scenarios of the system.
- These changing scenarios' effects on the system can be measured.
- The measurement has to be set against a benchmark.

Risk is never instantaneous. It has a time horizon. For instance, a firm engaged in a risky research and development venture will face significant amounts of risk but only until the product is fully developed or has proven itself in the market. These risks are caused by uncertainties in the technology of the product under research, uncertainties about the potential market, uncertainties about the level of competitive threats and substitutes, and so forth. These uncertainties will change over the course of the company's research and marketing activities—some uncertainties will increase while others will most likely decrease through the passage of time, actions, and events. However, only the uncertainties that affect the product directly will have any bearing on the risks of the product being successful. That is, only uncertainties that change the possible scenario outcomes will make the product risky (e.g., market and economic conditions). Finally, risk exists if it can be measured and compared against a benchmark. If no benchmark exists, then perhaps the conditions just described are the norm for research and development activities, and thus the negative results are to be expected. These benchmarks have to be measurable and tangible, for example, gross profits, success rates, market share, time to implementation, and so forth.

> Risk is any uncertainty that affects a system in an unknown fashion and its ramifications are unknown, but it brings great fluctuation in value and outcome. Risk has a time horizon, meaning that uncertainty evolves over time, which affects measurable future outcomes and scenarios with respect to a benchmark.

THE NATURE OF RISK AND RETURN

Nobel Laureate Harry Markowitz's groundbreaking research into the nature of risk and return has revolutionized the world of finance. His seminal work, which is now known all over the world as the *Markowitz Efficient Frontier*, looks at the nature of risk and return. Markowitz did not look at risk as the enemy but as a condition that should be embraced and balanced out through its expected returns. The concept of risk and return was then refined through later works by William Sharpe and others, who stated that a heightened risk necessitates a higher return, as elegantly expressed through the *capital asset pricing model* (CAPM), where the required rate of return on a marketable risky equity is equivalent to the return on an equivalent riskless asset plus a beta systematic and undiversifiable risk measure multiplied by the market risk's return premium. In essence, a higher risk asset requires a higher return. In Markowitz's model, one could strike a balance between risk and return. Depending on the risk appetite of an investor, the optimal or best-case returns can be obtained through the efficient frontier. Should the investor require a higher level of returns, he or she would have to face a higher level of risk. Markowitz's work carried over to finding combinations of individual projects or assets in a portfolio that would provide the best *bang for the buck*, striking an elegant balance between risk and return. In order to better understand this balance, also known as *risk adjustment* in modern risk analysis language, risks must first be measured and understood. The following section illustrates how risk can be measured.

THE STATISTICS OF RISK

The study of statistics refers to the collection, presentation, analysis, and utilization of numerical data to infer and make decisions in the face of uncertainty, where the actual population data is unknown. There are two branches in the study of statistics: descriptive statistics, where data is summarized and described, and inferential statistics, where the *population* is generalized through a small random sample, such that the *sample* becomes useful for making predictions or decisions when the population characteristics are unknown.

A sample can be defined as a subset of the population being measured, whereas the population can be defined as all possible observations of interest of a variable. For instance, if one is interested in the voting practices of all U.S. registered voters, the entire pool of a hundred million registered voters is considered the population, whereas a small survey of one thousand registered voters taken from several small towns across the nation is the sample. The calculated characteristics of the sample (e.g., mean, median, standard deviation) are termed *statistics*, while *parameters* imply that the entire population has been surveyed and the results tabulated. Thus, in decision making, the statistic is of vital importance, seeing that sometimes the entire population is yet unknown (e.g., who are all your customers, what is the total market share, etc.) or it is very difficult to obtain all relevant information on the population seeing that it would be too time- or resource-consuming.

In inferential statistics, the usual steps undertaken include:

- Designing the experiment—this phase includes designing the ways to collect all possible and relevant data.
- Collection of sample data—data is gathered and tabulated.
- Analysis of data—statistical analysis is performed.
- Estimation or prediction—inferences are made based on the statistics obtained.
- Hypothesis testing—decisions are tested against the data to see the outcomes.
- Goodness-of-fit—actual data is compared to historical data to see how accurate, valid, and reliable the inference is.
- Decision making—decisions are made based on the outcome of the inference.

Measuring the Center of the Distribution—The First Moment

The first moment of a distribution measures the *expected rate of return* on a particular project. It measures the location of the project's scenarios and possible outcomes on average. The common statistics for the first moment include the mean (average), median (center of a distribution), and mode (most commonly occurring value). Figure 2.1 illustrates the first moment—where, in this case, the first moment of this distribution is measured by the mean (μ) or average value.

Measuring the Spread of the Distribution—The Second Moment

The second moment measures the spread of a distribution, which is a *measure of risk*. The spread or width of a distribution measures the variability of

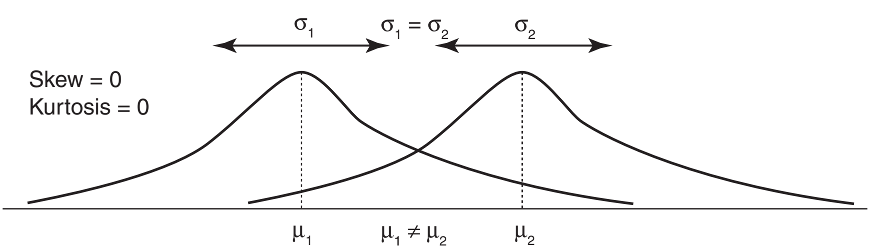


FIGURE 2.1 First moment.

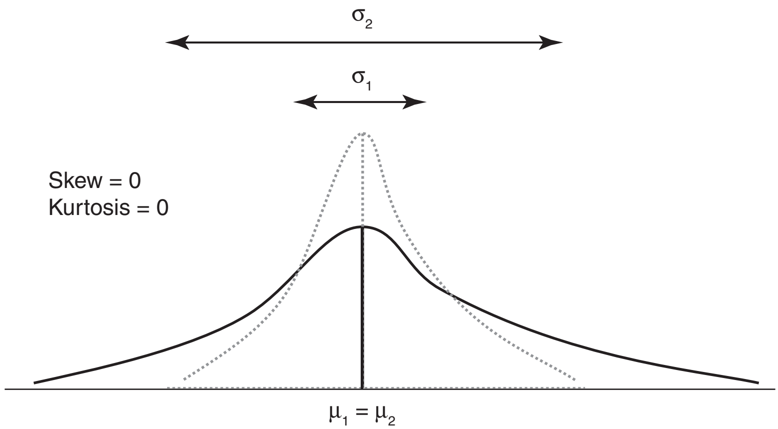


FIGURE 2.2 Second moment.

a variable, that is, the potential that the variable can fall into different regions of the distribution—in other words, the potential scenarios of outcomes. Figure 2.2 illustrates two distributions with identical first moments (identical means) but very different second moments or risks. The visualization becomes clearer in Figure 2.3. As an example, suppose there are two stocks and the first stock's movements (illustrated by the darker line) with the smaller fluctuation is compared against the second stock's movements (illustrated by the dotted line) with a much higher price fluctuation. Clearly an investor would view the stock with the wilder fluctuation as riskier because the outcomes of the more risky stock are relatively more unknown

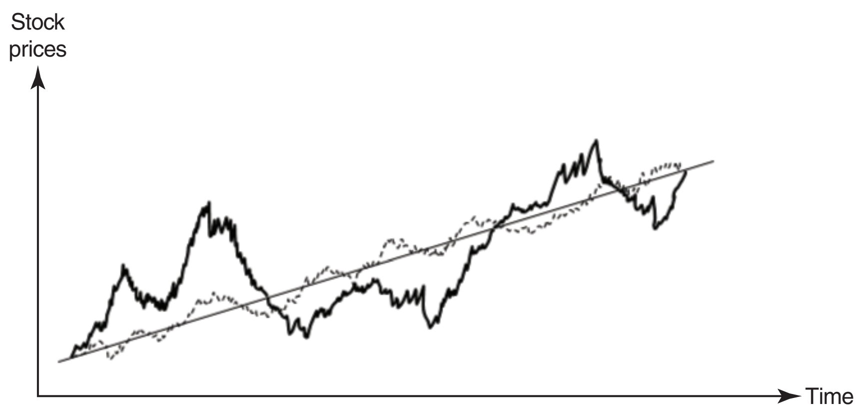


FIGURE 2.3 Stock price fluctuations.

than the less risky stock. The vertical axis in Figure 2.3 measures the stock prices; thus, the more risky stock has a wider range of potential outcomes. This range is translated into a distribution's width (the horizontal axis) in Figure 2.2, where the wider distribution represents the riskier asset. Hence, width or spread of a distribution measures a variable's risks.

Notice that in Figure 2.2, both distributions have identical first moments or central tendencies, but clearly the distributions are very different. This difference in the distributional width is measurable. Mathematically and statistically, the width or risk of a variable can be measured through several different statistics, including the range, standard deviation (σ), variance, coefficient of variation, volatility, and percentiles.

Measuring the Skew of the Distribution—The Third Moment

The third moment measures a distribution's skewness, that is, how the distribution is pulled to one side or the other. Figure 2.4 illustrates a negative or left skew (the tail of the distribution points to the left) and Figure 2.5 illustrates a positive or right skew (the tail of the distribution points to the right). The mean is always skewed toward the tail of the distribution while the median remains constant. Another way of seeing this is that the mean

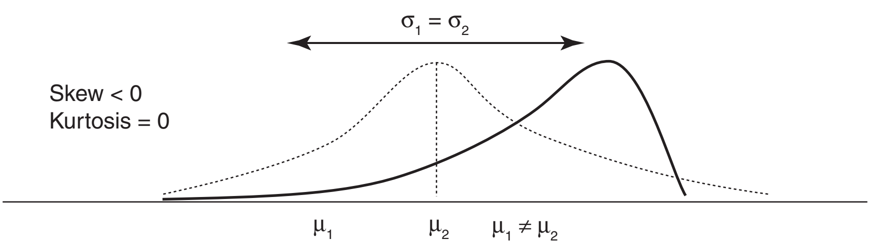


FIGURE 2.4 Third moment (left skew).

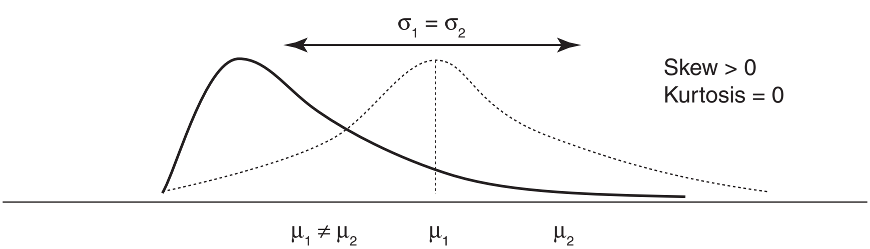


FIGURE 2.5 Third moment (right skew).

moves, but the standard deviation, variance, or width may still remain constant. If the third moment is not considered, then looking only at the expected returns (e.g., mean or median) and risk (standard deviation), a positively skewed project might be incorrectly chosen! For example, if the horizontal axis represents the net revenues of a project, then clearly a left or negatively skewed distribution might be preferred as there is a higher probability of greater returns (Figure 2.4) as compared to a higher probability for lower level returns (Figure 2.5). Thus, in a skewed distribution, the median is a better measure of returns, as the medians for both Figures 2.4 and 2.5 are identical, risks are identical, and, hence, a project with a negatively skewed distribution of net profits is a better choice. Failure to account for a project's distributional skewness may mean that the incorrect project may be chosen (e.g., two projects may have identical first and second moments, that is, they both have identical returns and risk profiles, but their distributional skews may be very different).

Measuring the Catastrophic Tail Events of the Distribution—The Fourth Moment

The fourth moment, or kurtosis, measures the peakedness of a distribution. Figure 2.6 illustrates this effect. The background (denoted by the dotted line) is a normal distribution with an excess kurtosis of 0. The new distribution has a higher kurtosis; thus the area under the curve is thicker at the tails with less area in the central body. This condition has major impacts on risk analysis as for the two distributions in Figure 2.6; the first three moments (mean, standard deviation, and skewness) can be identical, but the fourth moment (kurtosis) is different. This condition means that, although the returns and risks are identical, the probabilities of extreme and catastrophic

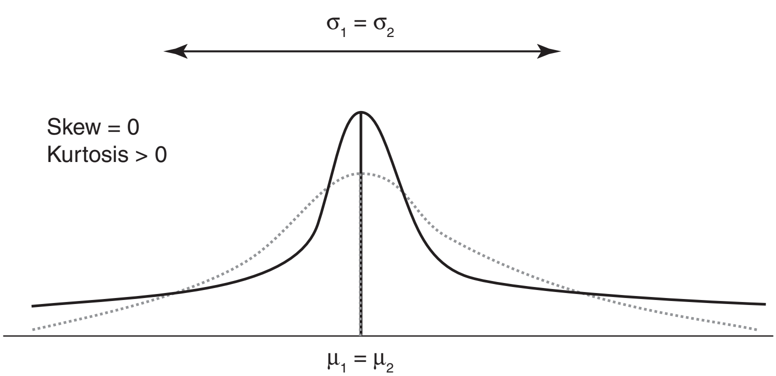


FIGURE 2.6 Fourth moment.

events (potential large losses or large gains) occurring are higher for a high kurtosis distribution (e.g., stock market returns are leptokurtic or have high kurtosis). Ignoring a project's return's kurtosis may be detrimental. Note that sometimes a normal kurtosis is denoted as 3.0, but in this book we use the measure of excess kurtosis, henceforth simply known as kurtosis. In other words, a kurtosis of 3.5 is also known as an excess kurtosis of 0.5, indicating that the distribution has 0.5 additional kurtosis above the normal distribution. The use of excess kurtosis is more prevalent in academic literature and is, hence, used here. Finally, the normalization of kurtosis to a base of 0 makes for easier interpretation of the statistic (e.g., a positive kurtosis indicates fatter-tailed distributions while negative kurtosis indicates thinner-tailed distributions).

> Most distributions can be defined up to four moments. The first moment describes the distribution's location or central tendency (expected returns), the second moment describes its width or spread (risks), the third moment its directional skew (most probable events), and the fourth moment its peakedness or thickness in the tails (catastrophic losses or gains). All four moments should be calculated and interpreted to provide a more comprehensive view of the project under analysis.

THE MEASUREMENTS OF RISK

There are multiple ways to measure risk in projects. This section summarizes some of the more common measures of risk and lists their potential benefits and pitfalls. The measures include:

- *Probability of Occurrence*. This approach is simplistic and yet effective. As an example, there is a 10 percent probability that a project will not break even (it will return a negative net present value indicating losses) within the next 5 years. Further, suppose two similar projects have identical implementation costs and expected returns. Based on a single-point estimate, management should be indifferent between them. However, if risk analysis such as Monte Carlo simulation is performed, the first project might reveal a 70 percent probability of losses compared to only a 5 percent probability of losses on the second project. Clearly, the second project is better when risks are analyzed.
- *Standard Deviation and Variance*. Standard deviation is a measure of the average of each data point's deviation from the mean.[2] This is the

most popular measure of risk, where a higher standard deviation implies a wider distributional width and, thus, carries a higher risk. The drawback of this measure is that both the upside and downside variations are included in the computation of the standard deviation. Some analysts define risks as the potential losses or downside; thus, standard deviation and variance will penalize upswings as well as downsides.

- *Semi-Standard Deviation.* The semi-standard deviation only measures the standard deviation of the downside risks and ignores the upside fluctuations. Modifications of the semi-standard deviation include calculating only the values below the mean, or values below a threshold (e.g., negative profits or negative cash flows). This provides a better picture of downside risk but is more difficult to estimate.
- *Volatility.* The concept of volatility is widely used in the applications of real options and can be defined briefly as a measure of uncertainty and risks.[3] Volatility can be estimated using multiple methods, including simulation of the uncertain variables impacting a particular project and estimating the standard deviation of the resulting asset's logarithmic returns over time. This concept is more difficult to define and estimate but is more powerful than most other risk measures in that this single value incorporates all sources of uncertainty rolled into one value.
- *Beta.* Beta is another common measure of risk in the investment finance arena. Beta can be defined simply as the undiversifiable, systematic risk of a financial asset. This concept is made famous through the CAPM, where a higher beta means a higher risk, which in turn requires a higher expected return on the asset.
- *Coefficient of Variation.* The coefficient of variation is simply defined as the ratio of standard deviation to the mean, which means that the risks are common-sized. For example, the distribution of a group of students' heights (measured in meters) can be compared to the distribution of the students' weights (measured in kilograms).[4] This measure of risk or dispersion is applicable when the variables' estimates, measures, magnitudes, or units differ.
- *Value at Risk.* Value at Risk (VaR) was made famous by J. P. Morgan in the mid-1990s through the introduction of its *RiskMetrics* approach, and has thus far been sanctioned by several bank governing bodies around the world. Briefly, it measures the amount of capital reserves at risk given a particular holding period at a particular probability of loss. This measurement can be modified to risk applications by stating, for example, the amount of potential losses a certain percent of the time during the period of the economic life of the project—clearly, a project with a smaller VaR is better.
- *Worst-Case Scenario and Regret.* Another simple measure is the value of the worst-case scenario given catastrophic losses. Another definition is

regret. That is, if a decision is made to pursue a particular project, but if the project becomes unprofitable and suffers a loss, the level of regret is simply the difference between the actual losses compared to doing nothing at all.

- *Risk-Adjusted Return on Capital.* Risk-adjusted return on capital (RAROC) takes the ratio of the difference between the fiftieth percentile (median) return and the fifth percentile return on a project to its standard deviation. This approach is used mostly by banks to estimate returns subject to their risks by measuring only the potential downside effects and ignoring the positive upswings.

The following appendix details the computations of some of these risk measures and is worthy of review before proceeding through the book.

APPENDIX—COMPUTING RISK

This appendix illustrates how some of the more common measures of risk are computed. Each risk measurement has its own computations and uses. For example, certain risk measures are applicable only on time-series data (e.g., volatility) while others are applicable in both cross-sectional and time-series data (e.g., variance, standard deviation, and covariance), while others require a consistent holding period (e.g., Value at Risk) or a market comparable or benchmark (e.g., beta coefficient).

Probability of Occurrence

This approach is simplistic yet effective. The probability of success or failure can be determined several ways. The first is through management expectations and assumptions, also known as expert opinion, based on historical occurrences or experience of the expert. Another approach is simply to gather available historical or comparable data, industry averages, academic research, or other third-party sources, indicating the historical probabilities of success or failure (e.g., pharmaceutical R&D's probability of technical success based on various drug indications can be obtained from external research consulting groups). Finally, Monte Carlo simulation can be run on a model with multiple interacting input assumptions and the output of interest (e.g., net present value, gross margin, tolerance ratios, and development success rates) can be captured as a simulation forecast and the relevant probabilities can be obtained, such as the probability of breaking even, probability of failure, probability of making a profit, and so forth. See Chapter 5 on step-by-step instructions on running and interpreting simulations and probabilities.

Standard Deviation and Variance

Standard deviation is a measure of the average of each data point's deviation from the mean. A higher standard deviation or variance implies a wider distributional width and, thus, a higher risk.

The standard deviation can be measured in terms of the population or sample, and for illustration purposes, is shown in the following list, where we define x_i as the individual data points, μ as the population mean, N as the population size, $\bar{x}$ as the sample mean, and n as the sample size:

Population standard deviation:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}\left(x_i - \mu\right)^2}{N}}$$

and population variance is simply the square of the standard deviation or σ^2. Alternatively, use Excel's *STDEVP* and *VARP* functions for the population standard deviation and variance respectively.

Sample standard deviation:

$$s = \sqrt{\frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}{n - 1}}$$

and sample variance is similarly the square of the standard deviation or s^2. Alternatively, use Excel's *STDEV* and *VAR* functions for the sample standard deviation and variance respectively. Figure 2.7 shows the step-by-step computations.

The drawbacks of this measure is that both the upside and downside variations are included in the computation of the standard deviation, and its dependence on the units (e.g., values of x in thousands of dollars versus millions of dollars are not comparable). Some analysts define risks as the potential losses or downside; thus, standard deviation and variance penalize upswings as well as downsides. An alternative is the semi-standard deviation.

Semi-Standard Deviation

The semi-standard deviation only measures the standard deviation of the downside risks and ignores the upside fluctuations. Modifications of the semi-standard deviation include calculating only the values below the mean, or values below a threshold (e.g., negative profits or negative cash flows). This

	X	X – Mean	Square of (X – Mean)
	–10.50	–9.07	82.2908
	12.25	13.68	187.1033
	–11.50	–10.07	101.4337
	13.25	14.68	215.4605
	–14.65	–13.22	174.8062
	15.65	17.08	291.6776
	–14.50	–13.07	170.8622
Sum	–10.00		
Mean	–1.43		

Population Standard Deviation and Variance

Sum of Square (X – Mean)	1223.6343
Variance = Sum of Square (X – Mean)/N	174.8049
Using Excel's VARP function:	174.8049
Standard Deviation = Square Root of (Sum of Square (X – Mean)/N)	13.2214
Using Excel's STDEVP function:	13.2214

Sample Standard Deviation and Variance

Sum of Square (X – Mean)	1223.6343
Variance = Sum of Square (X – Mean)/(N – 1)	203.9390
Using Excel's VAR function:	203.9390
Standard Deviation = Square Root of (Sum of Square (X – Mean)/(N–1))	14.2807
Using Excel's STDEV function:	14.2807

FIGURE 2.7 Standard deviation and variance computation.

approach provides a better picture of downside risk but is more difficult to estimate. Figure 2.8 shows how a sample semi-standard deviation and semi-variance are computed. Note that the computation must be performed manually.

Volatility

The concept of volatility is widely used in the applications of real options and can be defined briefly as a measure of uncertainty and risks. Volatility can be estimated using multiple methods, including simulation of the uncertain variables impacting a particular project and estimating the standard deviation of the resulting asset's logarithmic returns over time. This concept is more difficult to define and estimate but is more powerful than most other risk measures in that this single value incorporates all sources of uncertainty

	X	X – Mean	Square of (X – Mean)	
	–10.50	2.29	5.2327	
	12.25	Ignore		(Ignore the positive values)
	–11.50	1.29	1.6577	
	13.25	Ignore		(Ignore the positive values)
	–14.65	–1.86	3.4689	
	15.65	Ignore		(Ignore the positive values)
	–14.50	–1.71	2.9327	
Sum			–51.1500	
Mean			–12.7875	

Population Standard Deviation and Variance

Sum of Square (*X* – Mean)	13.2919
Variance = Sum of Square (*X* – Mean)/*N*	3.3230
Using Excel's VARP function:	3.3230
Standard Deviation = Square Root of (Sum of Square (*X* – Mean)/*N*)	1.8229
Using Excel's STDEVP function:	1.8229

Sample Standard Deviation and Variance

Sum of Square (*X* – Mean)	13.2919
Variance = Sum of Square (*X* – Mean)/(*N* – 1)	4.4306
Using Excel's VAR function:	4.4306
Standard Deviation = Square Root of (Sum of Square (*X* – Mean)/(*N*–1))	2.1049
Using Excel's STDEV function:	2.1049

FIGURE 2.8 Semi-standard deviation and semi-variance computation.

rolled into one value. Figure 2.9 illustrates the computation of an annualized volatility. Volatility is typically computed for time-series data only (i.e., data that follows a time series such as stock price, price of oil, interest rates, and so forth). The first step is to determine the relative returns from period to period, take their natural logarithms (ln), and then compute the sample standard deviation of these logged values. The result is the periodic volatility. Then, annualize the volatility by multiplying this periodic volatility by the square root of the number of periods in a year (e.g., 1 if annual data, 4 if quarterly data, and 12 if monthly data are used).

For a more detailed discussion of volatility computation as well as other methods for computing volatility such as using logarithmic present value approach, management assumptions, and GARCH, or generalized autoregressive conditional heteroskedasticity models, and how a discount rate can be determined from volatility, see *Real Options Analysis*, Second Edition, by Johnathan Mun (Wiley 2005).

Months	X	Relative Returns	LN (Relative Returns)	Square of (LN Relative Returns – Average)
0	10.50			
1	12.25	1.17	0.1542	0.0101
2	11.50	0.94	–0.0632	0.0137
3	13.25	1.15	0.1417	0.0077
4	14.65	1.11	0.1004	0.0022
5	15.65	1.07	0.0660	0.0001
6	14.50	0.93	–0.0763	0.0169
Sum			0.3228	
Average			0.0538	

Sample Standard Deviation and Variance

Sum of Square (LN Relative Returns – Average)	0.0507
Volatility = Square Root of (Sum of Square (LN Relative Returns – Average)/N – 1)	10.07%
Using Excel's STDEV function on LN(Relative Returns):	10.07%
Annualized Volatility (Periodic Volatility × Square Root (Periods in a Year))	34.89%

FIGURE 2.9 Volatility computation.

Beta

Beta is another common measure of risk in the investment finance arena. Beta can be defined simply as the undiversifiable, systematic risk of a financial asset. This concept is made famous through the CAPM, where a higher beta means a higher risk, which in turn requires a higher expected return on the asset. The beta coefficient measures the relative movements of one asset value to a comparable benchmark or market portfolio; that is, we define the beta coefficient as:

$$\beta = \frac{Cov(x,m)}{Var(m)} = \frac{\rho_{x,m}\sigma_x\sigma_m}{\sigma_m^2}$$

where $Cov(x,m)$ is the population covariance between the asset x and the market or comparable benchmark m, $Var(m)$ is the population variance of m, where both can be computed in Excel using the *COVAR* and *VARP* functions. The computed beta will be for the population. In contrast, the sample beta coefficient is computed using the correlation coefficient between x and m or $\rho_{x,m}$ and the sample standard deviations of x and m or using s_x and s_m instead of σ_x and σ_m.

A beta of 1.0 implies that the relative movements or risk of x is identical to the relative movements of the benchmark (see Example 1 in Figure 2.10

Example 1: Similar fluctuations with the market

Months	*X*	Market Comparable *M*
0	10.50	11.50
1	12.25	13.25
2	11.50	12.50
3	13.25	14.25
4	14.65	15.65
5	15.65	16.65
6	14.50	15.50

Sample Beta

Correlation between *X* and *M* using Excel's CORREL:	1.0000
Standard deviation of *X* using Excel's STDEV:	1.8654
Standard deviation of *M* using Excel's STDEV:	1.8654
Beta Coefficient	*1.0000*
(Correlation *X* and *M* * Stdev *X* * Stdev *M*)/ (Stdev *M* * Stdev *M*)	

Population Beta

Covariance population using Excel's COVAR:	2.9827
Variance of *M* using Excel's VARP:	2.9827
Population Beta	*1.0000*
(Covariance population (*X*, *M*)/Variance (*M*))	

Example 2: Half the fluctuations of the market

Months	*X*	Market Comparable *M*
0	10.50	21.00
1	12.25	24.50
2	11.50	23.00
3	13.25	26.50
4	14.65	29.30
5	15.65	31.30
6	14.50	29.00

Sample Beta

Correlation between *X* and *M* using Excel's CORREL:	1.0000
Standard deviation of *X* using Excel's STDEV:	1.8654
Standard deviation of *M* using Excel's STDEV:	3.7308
Beta Coefficient	*0.5000*
(Correlation *X* and *M* * Stdev *X* * Stdev *M*)/ (Stdev *M* * Stdev *M*)	

Population Beta

Covariance population using Excel's COVAR:	5.9653
Variance of *M* using Excel's VARP:	11.9306
Population Beta	*0.5000*
(Covariance population (*X*, *M*)/Variance (*M*))	

FIGURE 2.10 Beta coefficient computation.

where the asset x is simply one unit less than the market asset m, but they both fluctuate at the same levels). Similarly, a beta of 0.5 implies that the relative movements or risk of x is half of the relative movements of the benchmark (see Example 2 in Figure 2.10 where the asset x is simply half the market's fluctuations m). Therefore, beta is a powerful measure but requires a comparable to which to benchmark its fluctuations.

Coefficient of Variation

The coefficient of variation (CV) is simply defined as the ratio of standard deviation to the mean, which means that the risks are common sized. For example, a distribution of a group of students' heights (measured in meters) can be compared to the distribution of the students' weights (measured in kilograms). This measure of risk or dispersion is applicable when the variables' estimates, measures, magnitudes, or units differ. For example, in the computations in Figure 2.7, the CV for the population is –9.25 or –9.99 for the sample. The CV is useful as a measure of risk per unit of return, or when inverted, can be used as a measure of bang for the buck or returns per unit of risk. Thus, in portfolio optimization, one would be interested in minimizing the CV or maximizing the inverse of the CV.

Value at Risk

Value at Risk (*VaR*) measures the amount of capital reserves at risk given a particular holding period at a particular probability of loss. This measurement can be modified to risk applications by stating, for example, the amount of potential losses a certain percent of the time during the period of the economic life of the project—clearly, a project with a smaller VaR is better. VaR has a holding time period requirement, typically one year or one month. It also has a percentile requirement, for example, a 99.9 percent one-tail confidence. There are also modifications for daily risk measures such as DEaR or Daily Earnings at Risk. The VaR or DEaR can be determined very easily using Risk Simulator; that is, create your risk model, run a simulation, look at the forecast chart, and enter in 99.9 percent as the right-tail probability of the distribution or 0.01 percent as the left-tail probability of the distribution, then read the VaR or DEaR directly off the forecast chart.

Worst-Case Scenario and Regret

Another simple measure is the value of the worst-case scenario given catastrophic losses. An additional definition is regret; that is, if a decision is made to pursue a particular project, but if the project becomes unprofitable and suffers a loss, the level of regret is simply the difference between the actual losses compared to doing nothing at all. This analysis is very similar to the VaR but is not time dependent. For instance, a financial return on investment model can be created and a simulation is run. The 5 percent worst-case scenario can be read directly from the forecast chart in Risk Simulator.

Risk-Adjusted Return on Capital

Risk-adjusted return on capital (RAROC) takes the ratio of the difference between the fiftieth percentile P_{50} or its median return and the fifth percentile P_5 return on a project to its standard deviation σ, written as:

$$RAROC = \frac{P_{50} - P_5}{\sigma}$$

This approach is used mostly by banks to estimate returns subject to their risks by measuring only the potential downside effects and truncating the distribution to the worst-case 5 percent of the time, ignoring the positive upswings, while at the same time common sizing to the risk measure of standard deviation. Thus, RAROC can be seen as a measure that combines standard deviation, CV, semi-standard deviation, and worst-case scenario analysis. This measure is useful when applied with Monte Carlo simulation, where the percentiles and standard deviation measurements required can be obtained through the forecast chart's statistics view in Risk Simulator.

QUESTIONS

1. What is the efficient frontier and when is it used?
2. What are inferential statistics and what steps are required in making inferences?
3. When is using standard deviation less desirable than using semi-standard deviation as a measure of risk?
4. If comparing three projects with similar first, second, and fourth moments, would you prefer a project that has no skew, a positive skew, or a negative skew?
5. If comparing three projects with similar first to third moments, would you prefer a project that is leptokurtic (high kurtosis), mesokurtic (average kurtosis), or platykurtic (low kurtosis)? Explain your reasoning with respect to a distribution's tail area. Under what conditions would your answer change?
6. What are the differences and similarities between Value at Risk and worst-case scenario as a measure of risk?

CHAPTER 3

A Guide to Model-Building Etiquette

The first step in risk analysis is the creation of a model. A model can range from a simple three-line calculation in an Excel spreadsheet (e.g., A + B = C) to a highly complicated and oftentimes convoluted series of interconnected spreadsheets. Creating a proper model takes time, patience, strategy, and practice. Evaluating or learning a complicated model passed down to you that was previously created by another analyst may be rather cumbersome. Even the person who built the model revisits it weeks or months later and tries to remember what was created can sometimes find it challenging. It is indeed difficult to understand what the model originator was thinking of when the model was first built. As most readers of this book are Excel users, this chapter lists some model building blocks that every professional model builder should at least consider implementing in his or her Excel spreadsheets.

> As a rule of thumb, always remember to document the model; separate the inputs from the calculations and the results; protect the models against tampering; make the model user-friendly; track changes made in the model; automate the model whenever possible; and consider model aesthetics.

DOCUMENT THE MODEL

One of the major considerations in model building is its documentation. Although this step is often overlooked, it is crucial in order to allow continuity, survivorship, and knowledge transfer from one generation of model builders to the next. Inheriting a model that is not documented from a predecessor will only frustrate the new user. Some items to consider in model documentation include the following:

- *Strategize the Look and Feel of the Model.* Before the model is built, the overall structure of the model should be considered. This conceptualization includes how many sections the model will contain (e.g., each workbook file applies to a division; while each workbook has 10 worksheets representing each department in the division; and each worksheet has three sections, representing the revenues, costs, and miscellaneous items) as well as how each of these sections are related, linked, or replicated from one another.
- *Naming Conventions.* Each of these workbooks and worksheets should have a proper name. The recommended approach is simply to provide each workbook and worksheet a descriptive name. However, one should always consider brevity in the naming convention but yet provide sufficient description of the model. If multiple iterations of the model are required, especially when the model is created by several individuals over time, the date and version numbers should be part of the model's file name for proper archiving, backup, and identification purposes.
- *Executive Summary.* In the first section of the model, there should always be a welcome page with an executive summary of the model. The summary may include the file name, location on a shared drive, version of the model, developers of the model, and any other pertinent information, including instructions, assumptions, caveats, warnings, or suggestions on using the model.
- *File Properties.* Make full use of Excel's file properties (*File* | *Properties*). This simple action may make the difference between an orphaned model and a model that users will have more faith in as to how current or updated it is (Figure 3.1).
- *Document Changes and Tweaks.* If multiple developers work on the model, when the model is saved, the changes, tweaks, edits, and modifications should always be documented such that any past actions can be undone should it become necessary. This simple practice also provides a method to track the changes that have been made versus a list of bugs or development requirements.
- *Illustrate Formulas.* Consider illustrating and documenting the formulas used in the model, especially when complicated equations and calculations are required. Use Excel's Equation Editor to do this (*Insert* | *Object* | *Create New* | *Microsoft Equation*), but also remember to provide a reference for more advanced models.
- *Results Interpretation.* In the executive summary, on the reports or results summary page, include instructions on how the final analytical

FIGURE 3.1 Excel's file properties dialog box.

results should be interpreted, including what assumptions are used when building the model, any theory the results pertain to, any reference material detailing the technical aspects of the model, data sources, and any conjectures made to obtain certain input parameters.

- *Reporting Structure.* A good model should have a final report after the inputs have been entered and the analysis is performed. This report may be as simple as a printable results worksheet or as a more sophisticated macro that creates a new document (e.g., Risk Simulator has a reporting function that provides detailed analysis on the input parameters and output results).
- *Model Navigation.* Consider how a novice user will navigate between modules, worksheets, or input cells. One consideration is to include navigational capabilities in the model. These navigational capabilities range from a simple set of naming conventions (e.g., sheets in a workbook can be named "1. Input Data," "2. Analysis," and "3. Results") where the user can quickly and easily identify the relevant worksheets by their tab names (Figure 3.2), to more sophisticated methods. More sophisticated navigational methods include using hyperlinks and Visual Basic for Applications (VBA) code.

FIGURE 3.2 Worksheet tab names.

For instance, in order to create hyperlinks to other sheets from a main navigational sheet, click on *Insert* | *Hyperlink* | *Place in This Document* in Excel. Choose the relevant worksheet to link to within the workbook (Figure 3.3). Place all these links in the main navigational sheet and place only the relevant links in each sheet (e.g., only the main menu and Step 2 in the analysis are available in the Step 1 worksheet). These links can also be named as "next" or "previous," to further assist the user in navigating a large model. The second and more protracted approach is to use VBA codes to navigate the model. Refer to the appendix at the end of this chapter—A Primer on VBA Modeling and Writing Macros—for sample VBA codes used in said navigation and automation.

Document the model by strategizing the look and feel of the model, have an adequate naming convention, have an executive summary, include model property descriptions, indicate the changes and tweaks made, illustrate difficult formulas, document how to interpret results, provide a reporting structure, and make sure the model is easy to navigate.

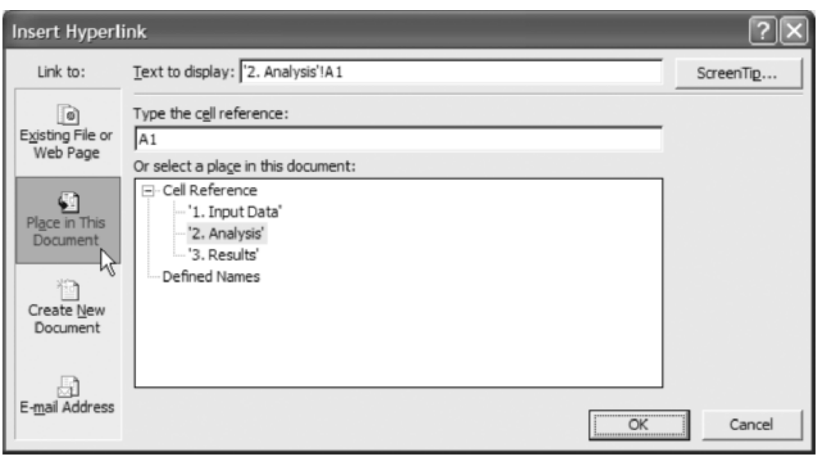


FIGURE 3.3 Insert hyperlink dialog box.

SEPARATE INPUTS, CALCULATIONS, AND RESULTS

- *Different Worksheets for Different Functions.* Consider using a different worksheet within a workbook for the model's input assumption (these assumptions should all be accumulated into a single sheet), a set of calculation worksheets, and a final set of worksheets summarizing the results. These sheets should then be appropriately named and grouped for easy identification. Sometimes, the input worksheet also has some key model results—this arrangement is very useful as a *management dashboard*, where slight tweaks and changes to the inputs can be made by management and the fluctuations in key results can be quickly viewed and captured.
- *Describe Input Variables.* In the input parameter worksheet, consider providing a summary of each input parameter, including where it is used in the model. Sometimes, this can be done through cell comments instead (*Insert | Comment*).
- *Name Input Parameter Cells.* Consider naming individual cells by selecting an input cell, typing the relevant name in the *Name Box* on the upper left corner of the spreadsheet, and hitting Enter (the arrow in Figure 3.4 shows the location of the name box). Also, consider naming ranges by selecting a range of cells and typing the relevant name in the *Name Box*. For more complicated models where multiple input parameters with similar functions exist, consider grouping these names. For instance, if the inputs "cost" and "revenues" exist in two different divisions, consider using the following hierarchical naming conventions (separated by periods in the names) for the Excel cells:

  Cost.Division.A
  Cost.Division.B
  Revenues.Division.A
  Revenues.Division.B

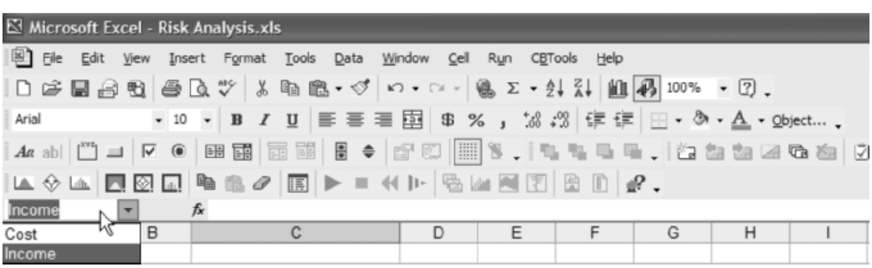


FIGURE 3.4 Name box in Excel.

- *Color Coding Inputs and Results.* Another form of identification is simply to color code the input cells one consistent color, while the results, which are usually mathematical functions based on the input assumptions and other intermediate calculations, should be color coded differently.
- *Model Growth and Modification.* A good model should always provide room for growth, enhancement, and update analysis over time. When additional divisions are added to the model, or other constraints and input assumptions are added at a later date, there should be room to maneuver. Another situation involves data updating, where, in the future, previous sales forecasts have now become reality and the actual sales now replace the forecasts. The model should be able to accommodate this situation. Providing the ability for data updating and model growth is where modeling strategy and experience count.
- *Report and Model Printing.* Always consider checking the overall model, results, summary, and report pages for their print layouts. Use Excel's *File* | *Print Preview* capability to set up the page appropriately for printing. Set up the headers and footers to reflect the dates of the analysis as well as the model version for easy comparison later. Use links, automatic fields, and formulas whenever appropriate (e.g., the Excel formula "*=Today()*" is a volatile field that updates automatically to the latest date when the spreadsheet model was last saved).

> Separate inputs, calculations, and results by creating different worksheets for different functions, describing input variables, naming input parameters, color coding inputs and results, providing room for model growth and subsequent modifications, and considering report and model printing layouts.

PROTECT THE MODELS

- *Protect Workbook and Worksheets.* Consider using spreadsheet protection (*Tools* | *Protection*) in your intermediate and final results summary sheet to prevent user tampering or accidental manipulation. Passwords are also recommended here for more sensitive models.[1]
- *Hiding and Protecting Formulas.* Consider setting cell properties to hide, lock, or both hide and lock cells (*Format* | *Cells* | *Protection*), then protect the worksheet (*Tools* | *Protection*) to prevent the user from accidentally overriding a formula (by locking a cell and protecting the

sheet), or still allow the user to see the formula without the ability to irreparably break the model by deleting the contents of a cell (by locking but not hiding the cell and protecting the sheet), or to prevent tampering with and viewing the formulas in the cell (by both locking and hiding the cell and then protecting the sheet).

Protect the models from user tampering at the workbook and worksheet levels through password protecting workbooks, or through hiding and protecting formulas in the individual worksheet cells.

MAKE THE MODEL USER-FRIENDLY: DATA VALIDATION AND ALERTS

- *Data Validation*. Consider preventing the user from entering bad inputs through spreadsheet validation. Prevent erroneous inputs through data validation (*Data* | *Validation* | *Settings*) where only specific inputs are allowed. Figure 3.5 illustrates data validation for a cell accepting only positive inputs. The *Edit* | *Copy* and *Edit* | *Paste Special* functions can be used to replicate the data validation if validation is chosen in the paste special command.
- *Error Alerts*. Provide error alerts to let the user know when an incorrect value is entered through data validation (*Data* | *Validation* | *Error Alert*)

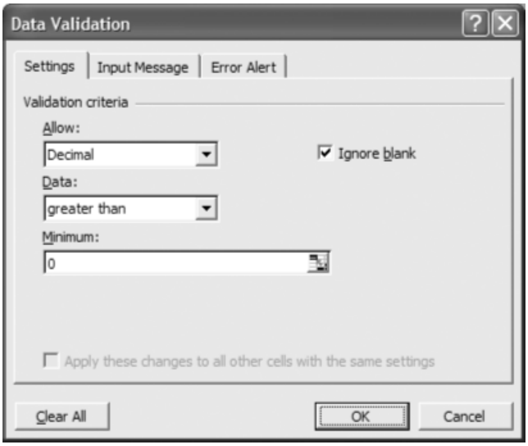


FIGURE 3.5 Data validation dialog box.

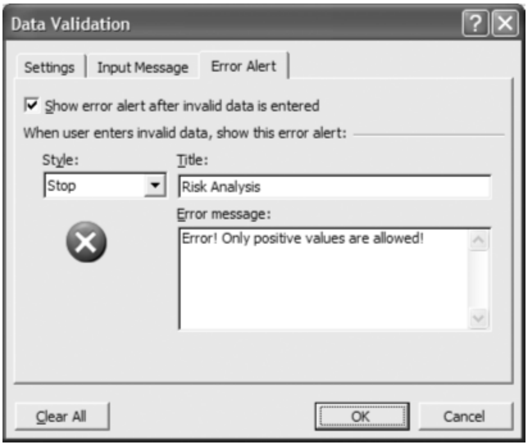


FIGURE 3.6 Error message setup for data validation.

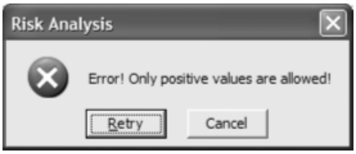


FIGURE 3.7 Error message for data validation.

shown in Figure 3.6. If the validation is violated, an error message box will be executed (Figure 3.7).

- *Cell Warnings and Input Messages*. Provide warnings and input messages when a cell is selected where the inputs required are ambiguous (*Data* | *Validation* | *Input Message*). The message box can be set up to appear whenever the cell is selected, regardless of the data validation. This message box can be used to provide additional information to the user about the specific input parameter or to provide suggested input values.
- *Define All Inputs*. Consider including a worksheet with named cells and ranges, complete with their respective definitions and where each variable is used in the model.

Make the model user-friendly through data validation, error alerts, cell warnings, and input messages, as well as defining all the inputs required in the model.

TRACK THE MODEL

- *Insert Comments*. Consider inserting comments for key variables (*Insert | Comment*) for easy recognition and for quick reference. Comments can be easily copied into different cells through the *Edit | Paste Special | Comments* procedure.
- *Track Changes*. Consider tracking changes if collaborating with other modelers (*Tools | Track Changes | Highlight Changes*). Tracking all changes is not only important, but it is also a courtesy to other model developers to note the changes and tweaks that were made.
- *Avoid Hard-Coding Values*. Consider using formulas whenever possible and avoid hard-coding numbers into cells other than assumptions and inputs. In complex models, it would be extremely difficult to track down where a model breaks because a few values are hard-coded instead of linked through equations. If a value needs to be hard-coded, it is by definition an input parameter and should be listed as such.
- *Use Linking and Embedding*. Consider object linking and embedding of files and objects (*Edit | Paste Special*) rather than using a simple paste function. This way, any changes in the source files can be reflected in the linked file. If linking between spreadsheets, Excel automatically updates these linked sheets every time the target sheet is opened. However, to avoid the irritating dialog pop-ups to update links every time the model is executed, simply turn off the warnings through *Edit | Links | Startup Prompt*.

> Track the model by inserting comments, using the track changes functionality, avoiding hard-coded values, and using the linking and embedding functionality.

AUTOMATE THE MODEL WITH VBA

Visual Basic for Applications is a powerful Excel tool that can assist in automating a significant amount of work. Although detailed VBA coding is beyond the scope of this book, an introduction to some VBA applications is provided in the appendix to this chapter—A Primer on VBA Modeling and Writing Macros—specifically addressing the following six automation issues:

1. Consider creating VBA modules for repetitive tasks (*Alt-F11* or *Tools | Macro | Visual Basic Editor*).

2. Add custom equations in place of complex and extended Excel equations.
3. Consider recording macros (*Tools* | *Macro* | *Record New Macro*) for repetitive tasks or calculations.
4. Consider placing automation forms in your model (*View* | *Toolbar* | *Forms*) and the relevant codes to support the desired actions.
5. Consider constraining users to only choosing specific inputs (*View* | *Toolbar* | *Forms*) and insert drop-list boxes and the relevant codes to support the desired actions.
6. Consider adding custom buttons and menu items on the user's model within Excel to locate and execute macros easily.

> Use VBA to automate the model, including adding custom equations, macros, automation forms, and predefined buttons.

MODEL AESTHETICS AND CONDITIONAL FORMATTING

- *Units.* Consider the input assumption's units and preset them accordingly in the cell to avoid any confusion. For instance, if a discount-rate input cell is required, the inputs can either be typed in as 20 or 0.2 to represent 20 percent. By avoiding a simple input ambiguity through preformatting the cells with the relevant units, user and model errors can be easily avoided.
- *Magnitude.* Consider the input's potential magnitude, where a large input value may obfuscate the cell's view by using the cell's default width. Change the format of the cell either to automatically reduce the font size to accommodate the higher magnitude input (*Format* | *Cells* | *Alignment* | *Shrink to Fit*) or have the cell width sufficiently large to accommodate all possible magnitudes of the input.
- *Text Wrapping and Zooming.* Consider wrapping long text in a cell (*Format* | *Cells* | *Alignment* | *Wrap Text*) for better aesthetics and view. This suggestion also applies to the zoom size of the spreadsheet. Remember that zoom size is worksheet specific and not workbook specific.
- *Merging Cells.* Consider merging cells in titles (*Format* | *Cells* | *Alignment* | *Merge Cells*) for a better look and feel.
- *Colors and Graphics.* Colors and graphics are an integral part of a model's aesthetics as well as a functional piece to determine if a cell is an input, a calculation, or a result. A careful blend of background colors and foreground graphics goes a long way in terms of model aesthetics.
- *Grouping.* Consider grouping repetitive columns or insignificant intermediate calculations (*Data* | *Group and Outline* | *Group*).

- *Hiding Rows and Columns.* Consider hiding extra rows and columns (select the relevant rows and columns to hide by selecting their row or column headers, and then choose *Format* | *Rows or Columns* | *Hide*) that are deemed as irrelevant intermediate calculations.
- *Conditional Formatting.* Consider conditional formatting such that if a cell's calculated result is a particular value (e.g., positive versus negative profits), the cell or font changes to a different color (*Format* | *Conditional Formatting*).
- *Auto Formatting.* Consider using Excel's auto formatting for tables (*Format* | *Auto Format*). Auto formatting will maintain the same look and feel throughout the entire Excel model for consistency.
- *Custom Styles.* The default Excel formatting can be easily altered, or alternatively, new styles can be added (*Format* | *Styles* | *New*). Styles can facilitate the model-building process in that consistent formatting is applied throughout the entire model by default and the modeler does not have to worry about specific cell formatting (e.g., shrink to fit and font size can be applied consistently throughout the model).
- *Custom Views.* In larger models where data inputs and output results are all over the place, consider using custom views (*View* | *Custom Views* | *Add*). This custom view feature allows the user to navigate through a large model spreadsheet with ease, especially when navigational macros are added to these views (see the appendix to this chapter—A Primer on VBA Modeling and Writing Macros—for navigating custom views using macros). In addition, different size zooms on areas of interest can be created within the same spreadsheet through custom views.

> Model aesthetics are preserved by considering the input units and magnitude, text wrapping and zooming views, cell merges, colors and graphics, grouping items, hiding excess rows and columns, conditional formatting, auto formatting, custom styles, and custom views.

APPENDIX—A PRIMER ON VBA MODELING AND WRITING MACROS

The Visual Basic Environment (VBE)

In Excel, access the VBE by hitting *Alt-F11* or *Tools* | *Macro* | *Visual Basic Environment*. The VBE looks like Figure 3.8. Select the VBA project pertaining to the opened Excel file (in this case, it is the *Risk Analysis.xls* file).

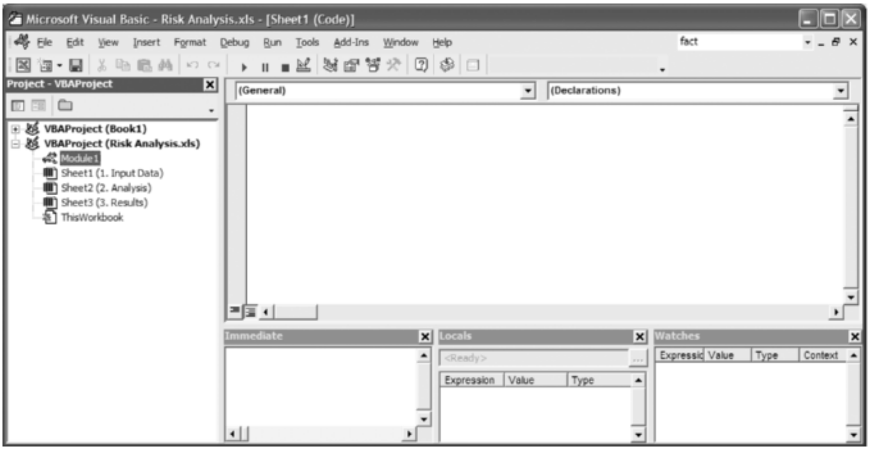


FIGURE 3.8 Visual basic environment.

Click on *Insert* | *Module* and double-click on the Module icon on the left window to open the module. You are now ready to start coding in VBA.

Custom Equations and Macros

Two Basic Equations The following example illustrates two basic equations. They are simple combination and permutation functions. Suppose that there are three variables, A, B, and C. Further suppose that two of these variables are chosen randomly. How many pairs of outcomes are possible? In a combination, order is not important and the following three pairs of outcomes are possible: AB, AC, and BC. In a permutation, order is important and matters; thus, the following six pairs of outcomes are possible: AB, AC, BA, BC, CA, and CB. The equations are:

$$Combination = \frac{(\text{Variable})!}{(\text{Choose})!(\text{Variable} - \text{Choose})!} = \frac{3!}{2!(3-2)!} = 3$$

$$Permutation = \frac{(\text{Variable})!}{(\text{Variable} - \text{Choose})!} = \frac{3!}{(3-2)!} = 6$$

If these two equations are widely used, then creating a VBA function will be more efficient and will avoid any unnecessary errors in larger models when Excel equations have to be created repeatedly. For instance, the manually inputted equation will have to be: *=fact(A1)/(fact(A2)*fact(A1–A2))* as compared to a custom function created in VBA where the function in Excel will now be =*combine*(*A1*,*A2*). The mathematical expression is exaggerated if the function is more complex, as will be seen later. The VBA code to be entered into the previous module (Figure 3.8) for the two simple equations is:

```
Public Function Combine(Variable As Double, Choose _
    As Double) As Double
Combine = Application.Fact(Variable) / (Application.Fact(Choose) * _
    Application.Fact(Variable – Choose))
End Function

Public Function Permute(Variable As Double, Choose As Double) _
    As Double
Permute = Application.Fact(Variable) / Application.Fact(Variable – _
    Choose)
End Function
```

Once the code is entered, the functions can be executed in the spreadsheet. The underscore at the end of a line of code indicates the continuation of the line of code on the next line.

Figure 3.9 shows the spreadsheet environment with the custom function. If multiple functions were entered, the user can also get access to those functions through the *Insert* | *Function* dialog wizard by choosing the user-defined category and scrolling down to the relevant functions (Figure 3.10). The functions arguments box comes up for the custom function chosen (Figure 3.11), and entering the relevant inputs or linking to input cells can be accomplished here.

Following are the VBA codes for the Black–Scholes models for estimating call and put options. The equations for the Black–Scholes are shown below and are simplified to functions in Excel named “BlackScholesCall” and “BlackScholesPut.”

$$Call = S\Phi\left[\frac{\ln(S / X) + (rf + \sigma^2 / 2)T}{\sigma\sqrt{T}}\right] - Xe^{-rf(T)}\Phi\left[\frac{\ln(S / X) + (rf - \sigma^2 / 2)T}{\sigma\sqrt{T}}\right]$$

$$Put = Xe^{-rf(T)}\Phi\left[-\frac{\ln(S / X) + (rf - \sigma^2 / 2)T}{\sigma\sqrt{T}}\right] - S\Phi\left[-\frac{\ln(S / X) + (rf + \sigma^2 / 2)T}{\sigma\sqrt{T}}\right]$$

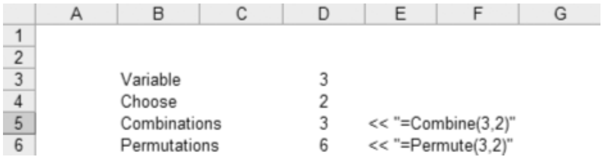


FIGURE 3.9 Excel spreadsheet with custom functions.

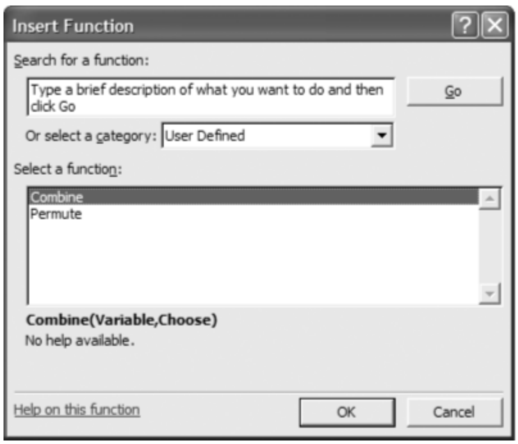


FIGURE 3.10 Insert function dialog box.

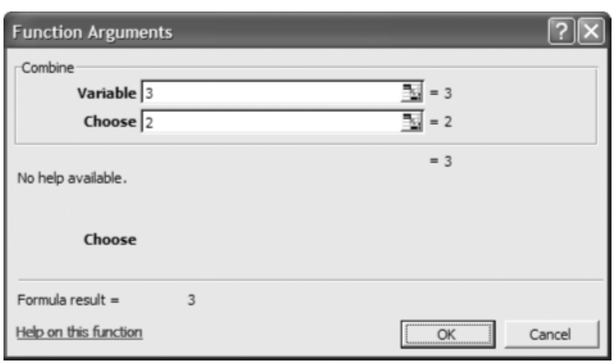


FIGURE 3.11 Function arguments box.

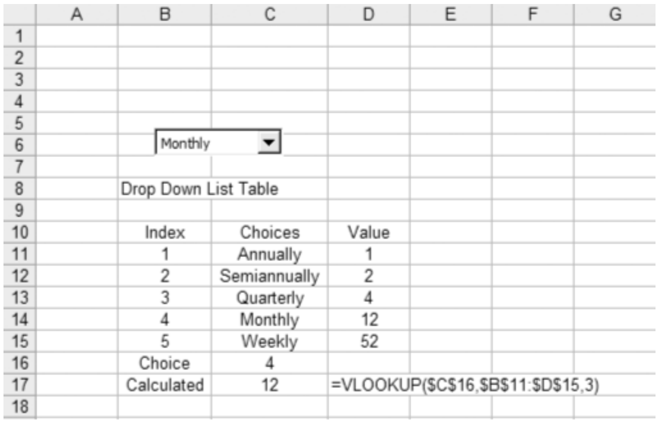
```
Public Function BlackScholesCall(Stock As Double, Strike As _
    Double, Time As Double, Riskfree _
    As Double, Volatility As Double) As Double
Dim D1 As Double, D2 As Double
D1 = (Log(Stock / Strike) + (Riskfree + 0.5 * Volatility ^ 2 / 2) * _
    Time) / (Volatility * Sqr(Time))
D2 = D1 – Volatility * Sqr(Time)
BlackScholesCall = Stock * Application.NormSDist(D1) – Strike * _
    Exp(–Time * Riskfree) * _
    Application.NormSDist(D2)
End Function

Public Function BlackScholesPut(Stock As Double, Strike As _
    Double Time As Double, Riskfree _
    As Double Volatility As Double) As Double
Dim D1 As Double, D2 As Double
D1 = (Log(Stock / Strike) + (Riskfree – 0.5 * Volatility ^ 2 / 2) * _
    Time) / (Volatility * Sqr(Time))
D2 = D1 – Volatility * Sqr(Time)
BlackScholesPut = Strike * Exp(–Time * Riskfree) * _
    Application.NormSDist(–D2) – Stock * _
    Application.NormSDist(–D1)
End Function
```

As an example, the function BlackScholesCall(100,100,1,5%,25%) results in 12.32 and BlackScholesPut(100,100,1,5%,25%) results in 7.44. Note that *Log* is a natural logarithm function in VBA and that *Sqr* is square root, and make sure there is a space before the underscore in the code. The underscore at the end of a line of code indicates the continuation of the line of code on the next line.

Form Macros

Another type of automation is form macros. In Excel, select *View* | *Toolbars* | *Forms* and the forms toolbar will appear. Click on the insert drop-list icon as shown in Figure 3.12 and drag it into an area in the spreadsheet to insert the drop list. Then create a drop-list table as seen in Figure 3.13 (cells B10 to D17). Point at the drop list and use the right mouse click to select *Format Control* | *Control*. Enter the input range as cells C11 to C15, cell link at C16, and five drop-down lines (Figure 3.14).

FIGURE 3.12 Forms icon bar.



	A	B	C	D	E	F	G
1							
2							
3							
4							
5							
6		Monthly					
7							
8		Drop Down List Table					
9							
10		Index	Choices	Value			
11		1	Annually	1			
12		2	Semiannually	2			
13		3	Quarterly	4			
14		4	Monthly	12			
15		5	Weekly	52			
16		Choice	4				
17		Calculated	12	=VLOOKUP($C$16,$B$11:$D$15,3)			
18							

FIGURE 3.13 Creating a drop-down box.

In Figure 3.13, the index column simply lists numbers 1 to n, where n is the total number of items in the drop-down list (in this example, n is 5). Here, the index simply converts the items (annually, semiannually, quarterly, monthly, and weekly) into corresponding indexes. The choices column in the input range is the named elements in the drop list. The value column lists the variables associated with the choice (semiannually means there are 2 periods in a year, or monthly means there are 12 periods in a year). Cell

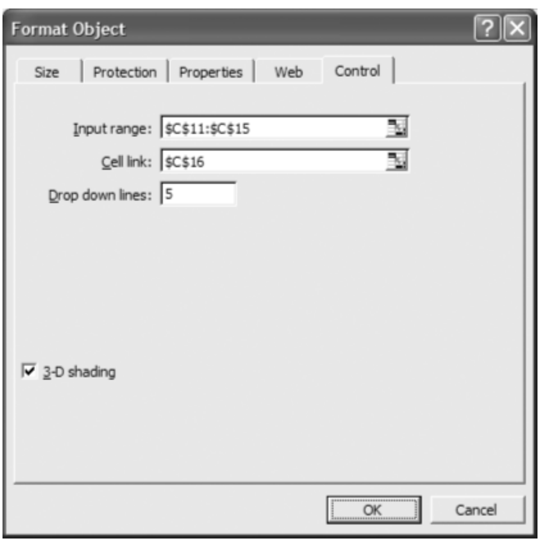


FIGURE 3.14 Format object dialog box.

C16 is the choice of the user selection; that is, if the user chooses monthly on the drop list, cell C16 will become 4, and so forth, as it is linked to the drop list in Figure 3.14. Cell C17 in Figure 3.13 is the equation

=*VLookup*($C$16,$B$11:$D$15, 3)

where the *VLookup* function will look up the value in cell C16 (the cell that changes in value depending on the drop-list item chosen) with respect to the first column in the area B11:D15, matches the corresponding row with the same value as in cell C16, and returns the value in the third column (3). In Figure 3.13, the value is 12. In other words, if the user chooses quarterly, then cell C16 will be 3, and cell C17 will be 4. Clearly, in proper model building, this entire table will be hidden somewhere out of the user's sight (placed in the extreme corners of the spreadsheet or in a distant corner and its font color changed to match the background, making it disappear or placed in a hidden worksheet). Only the drop list will be shown and the models will link to cell C17 as an input parameter. This situation forces the user to choose only from a list of predefined inputs and prevents any accidental insertion of invalid inputs.

Navigational VBA Codes A simple macro to navigate to sheet "2. Analysis" is shown here. This macro can be written in the VBA environment or recorded in the *Tools* | *Macros* | *Record New Macro*, then perform the relevant navigational actions (i.e., clicking on the "2. Analysis" sheet and hitting the stop recording button), return to the VBA environment, and open up the newly recorded macro.

```
Sub MoveToSheet2()
Sheets("2. Analysis").Select
End Sub
```

However, if custom views (*View* | *Custom Views* | *Add*) are created in Excel worksheets (to facilitate finding or viewing certain parts of the model such as inputs, outputs, etc.), navigations can also be created through the following, where a custom view named "results" had been previously created:

```
Sub CustomView()
ActiveWorkbook.CustomViews("Results").Show
End Sub
```

Form buttons can then be created and these navigational codes can be attached to the buttons. For instance, click on the fourth icon in the forms icon bar (Figure 3.12) and insert a form button in the spreadsheet and assign

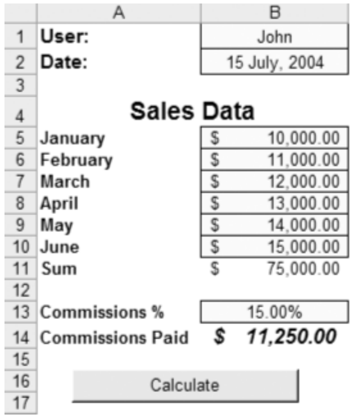


FIGURE 3.15 Simple automated model.

the relevant macros created previously. (If the select macro dialog does not appear, right-click the form button and select *Assign Macro*.)

Input Boxes Input boxes are also recommended for their ease of use. The following illustrates some sample input boxes created in VBA, where the user is prompted to enter certain restrictive inputs in different steps or wizards. For instance, Figure 3.15 illustrates a simple sales commission calculation model, where the user inputs are the colored and boxed cells. The resulting commissions (cell B11 times cell B13) will be calculated in cell B14. The user would start using the model by clicking on the *Calculate* form button. A series of input prompts will then walk the user through inputting the relevant assumptions (Figure 3.16).

The code can also be set up to check for relevant inputs, that is, sales commissions have to be between 0.01 and 0.99. The full VBA code is shown next. The code is first written in VBA, and then the form button is placed in the worksheet that calls the VBA code.

FIGURE 3.16 Sample input box.

```
Sub UserInputs()
Dim User As Variant, Today As String, Sales As Double, _
     Commissions As Double
Range("B1").Select
User = InputBox("Enter your name:")
ActiveCell.FormulaR1C1 = User
Range("B2").Select
Today = InputBox("Enter today's date:")
ActiveCell.FormulaR1C1 = Today

Range("B5").Select
Sales = InputBox("Enter the sales revenue:")
ActiveCell.FormulaR1C1 = Sales
Dim N As Double

For N = 1 To 5
   ActiveCell.Offset(1, 0).Select
   Sales = InputBox("Enter the sales revenue for the following _
        period:")
   ActiveCell.FormulaR1C1 = Sales
Next N

Range("B13").Select
Commissions = 0
Do While Commissions < 0.01 Or Commissions > 0.99
   Commissions = InputBox("Enter recommended commission rate _
        between 1% and 99%:")
Loop
ActiveCell.FormulaR1C1 = Commissions
Range("B1").Select
End Sub
```

Forms and Icons Sometimes, for globally used macros and VBA scripts, a menu item or an icon can be added to the user's spreadsheet. Insert a new menu item by clicking on *Tools* | *Customize* | *Commands* | *New Menu* and dragging the *New Menu* item list to the Excel menu bar to a location right before the *Help* menu. Click on *Modify Selection* and rename the menu item accordingly (e.g., Risk Analysis). Also, an ampersand ("&") can be placed before a letter in the menu item name to underline the next letter such that the menu can be accessed through the keyboard by hitting the *Alternate* key and then the corresponding letter key. Next, click on *Modify Selection* | *Begin a Group* and then drag the *New Menu* item list again to the menu bar,

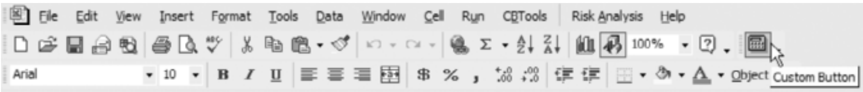


FIGURE 3.17 Custom menu and icon.

but this time, right under the Risk Analysis group. Now, select this submenu item and click on *Modify Selection* | *Name* and rename it *Run Commissions*. Then, *Modify Selection* | *Assign Macro* and assign it to the User Input macro created previously.

Another method to access macros (other than using menu items or *Tools* | *Macro* | *Macros*, or *Alt-F8*) is to create an icon on the icon toolbar. To do this, click on *Tools* | *Customize* | *Toolbars* | *New*. Name the new toolbar accordingly and drag it to its new location anywhere on the icon bar. Then, select the *Commands* | *Macros* | *Custom Button*. Drag the custom button icon to the new toolbar location. Select the new icon on the toolbar and click on *Modify Selection* | *Assign Macro*. Assign the User Input macro created previously. The default button image can also be changed by clicking on *Modify Selection* | *Change Button Image* and selecting the relevant icon accordingly, or from an external image file. Figure 3.17 illustrates the new menu item (Risk Analysis) and the new icon in the shape of a calculator, where selecting either the menu item or the icon will evoke the User Input macro, which walks the user through the simple input wizard.

EXERCISES

1. Create an Excel worksheet with each of the following components activated:
   a. Cells in an Excel spreadsheet with the following data validations: no negative numbers are allowed, only positive integers are allowed, only numerical values are allowed.
   b. Create a form macro drop list (see the appendix to this chapter) with the following 12 items in the drop list: January, February, March, . . . December. Make sure the selection of any item in the drop list will change a corresponding cell's value.
2. Go through the VBA examples in the appendix to this chapter and re-create the following macros and functions for use in an Excel spreadsheet:
   a. Create a column of future sales with the following equation for future sales (Years 2 to 11): *Future sales = (1+RAND())*(Past Year Sales)*

for 11 future periods starting with the current year's sales of $100 (Year 1). Then, in VBA, create a macro using the *For . . . Next* loop to simulate this calculation 1,000 times and insert a form button to activate the macro in the Excel worksheet.

b. Create the following income function in VBA for use in the Excel spreadsheet: *Income* = *Benefits* – *Cost*. Try out different benefits and cost inputs to make sure the function works properly.

PART Three

Risk Quantification

CHAPTER 4

On the Shores of Monaco

Monte Carlo simulation, named for the famous gambling capital of Monaco, is a very potent methodology. For the practitioner, simulation opens the door for solving difficult and complex but practical problems with great ease. Perhaps the most famous early use of Monte Carlo simulation was by the Nobel physicist Enrico Fermi (sometimes referred to as the father of the atomic bomb) in 1930, when he used a random method to calculate the properties of the newly discovered neutron. Monte Carlo methods were central to the simulations required for the Manhattan Project, where in the 1950s Monte Carlo simulation was used at Los Alamos for early work relating to the development of the hydrogen bomb, and became popularized in the fields of physics and operations research. The Rand Corporation and the U.S. Air Force were two of the major organizations responsible for funding and disseminating information on Monte Carlo methods during this time, and today there is a wide application of Monte Carlo simulation in many different fields including engineering, physics, research and development, business, and finance.

Simplistically, Monte Carlo simulation creates artificial futures by generating thousands and even hundreds of thousands of sample paths of outcomes and analyzes their prevalent characteristics. In practice, Monte Carlo simulation methods are used for risk analysis, risk quantification, sensitivity analysis, and prediction. An alternative to simulation is the use of highly complex stochastic closed-form mathematical models. For analysts in a company, taking graduate-level advanced math and statistics courses is just not logical or practical. A brilliant analyst would use all available tools at his or her disposal to obtain the same answer the easiest and most practical way possible. And in all cases, when modeled correctly, Monte Carlo simulation provides similar answers to the more mathematically elegant methods. In addition, there are many real-life applications where closed-form models do not exist and the only recourse is to apply simulation methods. So, what exactly is Monte Carlo simulation and how does it work?

WHAT IS MONTE CARLO SIMULATION?

Today, fast computers have made possible many complex computations that were seemingly intractable in past years. For scientists, engineers, statisticians, managers, business analysts, and others, computers have made it possible to create models that simulate reality and aid in making predictions, one of which is used in simulating real systems by accounting for randomness and future uncertainties through investigating hundreds and even thousands of different scenarios. The results are then compiled and used to make decisions. This is what Monte Carlo simulation is all about.

Monte Carlo simulation in its simplest form is a random number generator that is useful for forecasting, estimation, and risk analysis. A simulation calculates numerous scenarios of a model by repeatedly picking values from a user-predefined *probability distribution* for the uncertain variables and using those values for the model. As all those scenarios produce associated results in a model, each scenario can have a forecast. Forecasts are events (usually with formulas or functions) that you define as important outputs of the model.

Think of the Monte Carlo simulation approach as picking golf balls out of a large basket repeatedly with replacement. The size and shape of the basket depend on the distributional *input assumption* (e.g., a normal distribution with a mean of 100 and a standard deviation of 10, versus a uniform distribution or a triangular distribution) where some baskets are deeper or more symmetrical than others, allowing certain balls to be pulled out more frequently than others. The number of balls pulled repeatedly depends on the number of *trials* simulated. For a large model with multiple related assumptions, imagine the large model as a very large basket, where many baby baskets reside. Each baby basket has its own set of colored golf balls that are bouncing around. Sometimes these baby baskets are linked with each other (if there is a *correlation* between the variables), forcing the golf balls to bounce in tandem, whereas in other uncorrelated cases, the balls are bouncing independently of one another. The balls that are picked each time from these interactions within the model (the large basket) are tabulated and recorded, providing a *forecast output* result of the simulation.

WHY ARE SIMULATIONS IMPORTANT?

An example of why simulation is important can be seen in the case illustration in Figures 4.1 and 4.2, termed the Flaw of Averages.[1] The example is most certainly worthy of more detailed study. It shows how an analyst may be misled into making the wrong decisions without the use of simulation. Suppose you are the owner of a shop that sells perishable goods and you need to make a decision on the optimal inventory to have on hand. Your

Actual	5	Average	5.00
Inventory Held	6		
		Historical Data	(5 Yr)
Perishable Cost	$100	Month	Actual
Fed Ex Cost	$175	1	12
		2	11
Total Cost	$100	3	7
		4	0
		5	0
		6	2
		7	7
		8	0
		9	11
		10	12
		11	0
		12	9
		13	3
		14	5
		15	0
		16	2
		17	1
		18	10
		58	3
		59	2
		60	17

Your company is a retailer in perishable goods and you were tasked with finding the optimal level of inventory to have on hand. If your inventory exceeds actual demand, there is a $100 perishable cost while a $175 Fed Ex cost is incurred if your inventory is insufficient to cover the actual level of demand. These costs are on a per unit basis. Your first inclination is to collect historical demand data as seen on the right, for the past 60 months. You then take a simple average, which was found to be 5 units. Hence, you select 5 units as the optimal inventory level. You have just committed a major mistake called the Flaw of Averages!

The actual demand data are shown here on the right. Rows 19 through 57 are hidden to conserve space. Being the analyst, what must you then do?

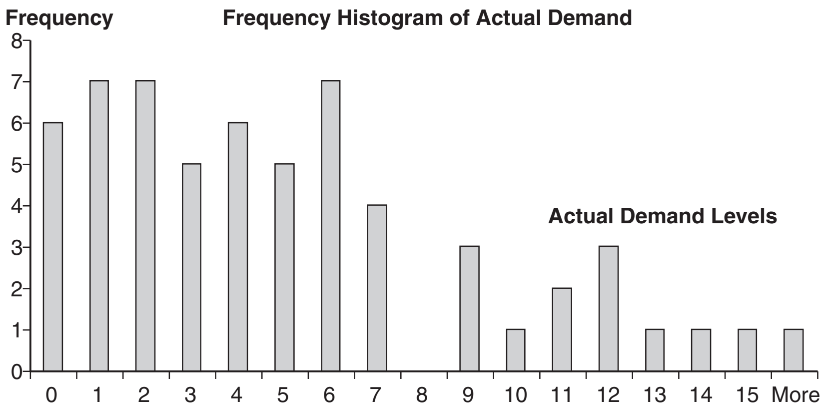


FIGURE 4.1 The flaw of averages example.

new-hire analyst was successful in downloading 5 years worth of monthly historical sales levels and she estimates the average to be five units. You then make the decision that the optimal inventory to have on hand is five units. You have just committed the flaw of averages. As the example shows, the

Simulated Average			
Actual Demand	8.53	Simulated Demand Range	From 7.21 and 9.85
Inventory Held	9.00	Simulated Cost Range	From 178.91 to 149

Perishable Cost	$100
Fed Ex Cost	$175
Total Cost	$46.88

The best method is to perform a nonparametric simulation where we use the actual historical demand levels as inputs to simulate the most probable level of demand going forward, which we found as 8.53 units. Given this demand, the lowest cost is obtained through a trial inventory of 9 units, a far cry from the original Flaw of Averages estimate of 5 units.

Trial Inventory	Total Cost
1.00	$1,318
2.00	$1,143
3.00	$968
4.00	$793
5.00	$618
6.00	$443
7.00	$268
8.00	$93
9.00	$47
10.00	$147
11.00	$247
12.00	$347
13.00	$447
14.00	$547
15.00	$647
16.00	$747

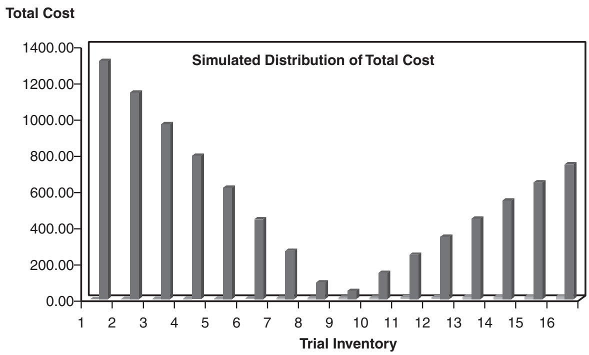


FIGURE 4.2 Fixing the flaw of averages with simulation.

obvious reason why this error occurs is that the distribution of historical demand is highly skewed while the cost structure is asymmetrical. For example, suppose you are in a meeting, and your boss asks what everyone made last year. You take a quick poll and realize that the salaries range from $60,000 to $150,000. You perform a quick calculation and find the average to be $100,000. Then, your boss tells you that he made $20 million last year! Suddenly, the average for the group becomes $1.5 million. This value of $1.5 million clearly in no way represents how much each of your peers made last year. In this case, the median may be more appropriate. Here you see that simply using the average will provide highly misleading results.[2]

Continuing with the example, Figure 4.2 shows how the right inventory level is calculated using simulation. The approach used here is called *nonparametric* bootstrap simulation. It is nonparametric because in this simulation approach, no distributional parameters are assigned. Instead of assuming some preset distribution (normal, triangular, lognormal, or the like) and its required parameters (mean, standard deviation, and so forth) as required in a Monte Carlo *parametric* simulation, nonparametric simulation uses the data themselves to tell the story.

Imagine that you collect 5 years worth of historical demand levels and write down the demand quantity on a golf ball for each month. Throw all 60 golf balls into a large basket and mix the basket randomly. Pick a golf ball out at random and write down its value on a piece of paper, then replace the ball in the basket and mix the basket again. Do this 60 times and calculate the average. This process is a single grouped trial. Perform this entire process several thousand times, with replacement. The distribution of these thousands of averages represents the outcome of the simulation forecast. The expected value of the simulation is simply the average value of these thousands of averages. Figure 4.2 shows an example of the distribution stemming from a nonparametric simulation. As you can see, the optimal inventory rate that minimizes carrying costs is nine units, far from the average value of five units previously calculated in Figure 4.1.

Clearly, each approach has its merits and disadvantages. Nonparametric simulation, which can be easily applied using Risk Simulator's custom distribution,[3] uses historical data to tell the story and to predict the future. Parametric simulation, however, forces the simulated outcomes to follow well-behaving distributions, which is desirable in most cases. Instead of having to worry about cleaning up any messy data (e.g., outliers and nonsensical values) as is required for nonparametric simulation, parametric simulation starts fresh every time.

> Monte Carlo simulation is a type of parametric simulation, where specific distributional parameters are required before a simulation can begin. The alternative approach is nonparametric simulation where the raw historical data is used to tell the story and no distributional parameters are required for the simulation to run.

COMPARING SIMULATION WITH TRADITIONAL ANALYSES

Figure 4.3 illustrates some traditional approaches used to deal with uncertainty and risk. The methods include performing sensitivity analysis, scenario analysis, and probabilistic scenarios. The next step is the application of Monte Carlo simulation, which can be seen as an extension to the next step in uncertainty and risk analysis. Figure 4.4 shows a more advanced use

(*Text continues on page 82.*)

Unit Sales	10	
Unit Price	$10	10 units × $10 per unit
Total Revenue	$100	
Unit Variable Cost	$5	
Fixed Cost	$20	$20 Fixed + ($5 × 10) Variable
Total Cost	$70	
Net Income	$30	$100 – $70

Point Estimates

This is a simple example of a Point Estimate approach. The issues that arise may include the risk of how confident you are in the unit sales projections, the sales price, and the variable unit cost.

Since the bottom line Net Income is the key financial performance indicator here, an uncertainty in future sales volume will be impounded into the Net Income calculation. How much faith do you have in your calculation based on a simple point estimate?

Recall the Flaw of Average example where a simple point estimate could yield disastrous conclusions.

Sensitivity Analysis

Here, we can make unit changes to the variables in our simple model to see the final effects of such a change. Looking at the simple example, we know that only Unit Sales, Unit Price, and Unit Variable Cost can change. This is since Total Revenues, Total Costs, and Net Income are calculated values while Fixed Cost is assumed to be fixed and unchanging, regardless of the amount of sales units or sales price. Changing these three variables by one unit shows that from the original $40, Net Income has now increased $5 for Unit Sales, increased $10 for Unit Price, and decreased $10 for Unit Variable Cost.

Unit Sales	11	Change 1 unit	Unit Sales	10		Unit Sales	10
Unit Price	$10		Unit Price	$11	Change 1 unit	Unit Price	$10
Total Revenue	$110		Total Revenue	$110		Total Revenue	$100
Unit Variable Cost	$5		Unit Variable Cost	$5	Change 1 unit	Unit Variable Cost	$6
Fixed Cost	$20	Up $10	Fixed Cost	$20		Fixed Cost	$20
Total Cost	$75	Up $5	Total Cost	$70		Total Cost	$80
Net Income	$35		Net Income	$40	Down $10	Net Income	$20

Hence, we know that Unit Price has the most positive impact on the Net Income bottom line and Unit Variable Cost the most negative impact. In terms of making assumptions, we know that additional care must be taken when forecasting and estimating these variables. However, we still are in the dark concerning which sensitivity set of results we should be looking at or using.

Scenario Analysis

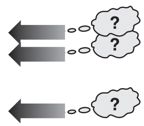
In order to provide an added element of variability, using the simple example above, you can perform a Scenario Analysis, where you would change values of key variables by certain units given certain assumed scenarios. For instance, you may assume three economic scenarios where unit sales and unit sale prices will vary. Under a good economic condition, unit sales go up to 14 at $11 per unit. Under a nominal economic scenario, units sales will be 10 units at $10 per unit. Under a bleak economic scenario, unit sales decrease to 8 units but prices per unit stays at $10.

Unit Sales	14	Good Economy	Unit Sales	10	Average Economy	Unit Sales	8
Unit Price	$11		Unit Price	$10		Unit Price	$10
Total Revenue	$154		Total Revenue	$100		Total Revenue	$80
Unit Variable Cost	$5		Unit Variable Cost	$5		Unit Variable Cost	$5
Fixed Cost	$20		Fixed Cost	$20	Bad Economy	Fixed Cost	$20
Total Cost	$90		Total Cost	$70		Total Cost	$60
Net Income	$64		Net Income	$30		Net Income	$20

Looking at the Net Income results, we have $64, $30 and $20. The problem here is, the variation is too large. Which condition do I think will most likely occur and which result do I use in my budget forecast for the firm? Although Scenario Analysis is useful in ascertaining the impact of different conditions, both advantageous and adverse, the analysis provides little insight to which result to use.

	Probability	Net Income
Good Economy	35%	$64.00
Average Economy	40%	$30.00
Bad Economy	25%	$20.00
EMV		$39.40

Unit Sales	10
Unit Price	$10
Total Revenue	$100
Unit Variable Cost	$5
Fixed Cost	$20
Total Cost	$70
Net Income	$30



By performing the simulation thousands of times, we essentially perform thousands of sensitivity analyses and scenario analyses given different sets of probabilities. These are all set in the original simulation assumptions (types of probability distributions, the parameters of the distributions and which variables to simulate).

The results calculated from the simulation output can then be interpreted as follows:

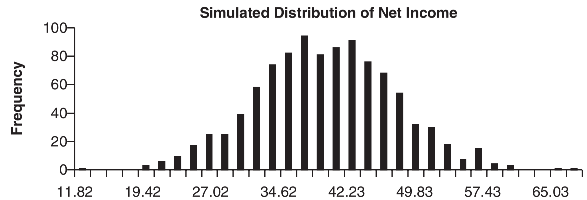
Discussions about types of distributional assumptions to use and the actual simulation approach will be discussed later.

Probabilistic Scenario Analysis

We can always assign probabilities that each scenario will occur, creating a Probabilistic Scenario Analysis and simply calculate the Expected Monetary Value (EMV) of the forecasts. The results here are more robust and reliable than a simple scenario analysis since we have collapsed the entire range of potential outcomes of $64, $30, and $20 into a single expected value. This value is what you would expect to get on average.

Simulation Analysis

Looking at the original model, we know that through Sensitivity Analysis, Unit Sales, Unit Price and Unit Variable Cost are three highly uncertain variables. We can then very easily simulate these three unknowns thousands of times (based on certain distributional assumptions) to see what the final Net Income value looks like.



Average	$40.04
Median	$39.98
Mode	$46.63
Standard Deviation	$8.20
95% Confidence Interval	Between $56.16 and $24.09

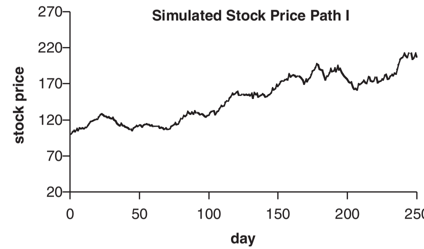
FIGURE 4.3 Point estimates, sensitivity analysis, scenario analysis, and simulation.

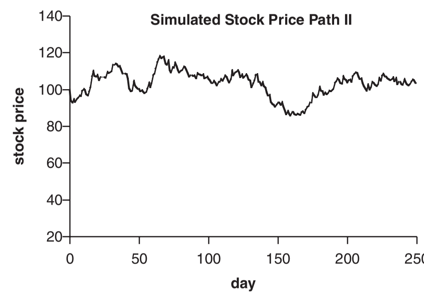
A Simple Simulation Example
We need to perform many simulations to obtain a valid distribution.

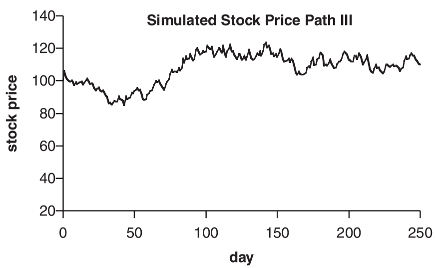
Mean	15%
Sigma	30%
Timing	Daily
Starting Value	100

Here we see the effects of performing a simulation of stock price paths following a Geometric Brownian Motion model for daily closing prices. Three sample paths are seen here. In reality, thousands of simulations are performed and their distributional properties are analyzed. Frequently, the average closing prices of these thousands of simulations are analyzed, based on these simulated price paths.

time days	normal deviates	value simulated
		100.0000
1	0.0873	100.2259
2	–0.4320	99.4675
3	–0.1389	99.2652
4	–0.4583	98.4649
5	1.7807	101.9095
6	–1.4406	99.2212
7	–0.5577	98.2357
8	0.5277	99.2838
9	–0.4844	98.4345
10	–0.2307	98.0634
11	0.8688	99.7532
12	2.1195	83.9088
13	–1.9756	100.1461
14	1.3734	102.8517
15	–0.8790	101.2112
16	–0.7610	99.8203
17	0.3168	100.4824
18	–0.0511	100.4452
19	0.0653	100.6301
20	–0.6073	99.5368
21	0.6900	100.9091
22	–0.7012	99.6353
23	1.4784	102.5312
24	–0.9195	100.8184
25	–0.3343	100.2411
26	–2.3395	95.9465
27	–1.7831	92.8103
28	–0.3247	92.2958
29	0.5053	93.2409
30	0.0386	93.3652
247	1.0418	100.9205
248	–0.7052	99.6388
249	0.1338	99.9521
250	0.0451	100.0978







Rows 31 through 246 have been hidden to conserve space.

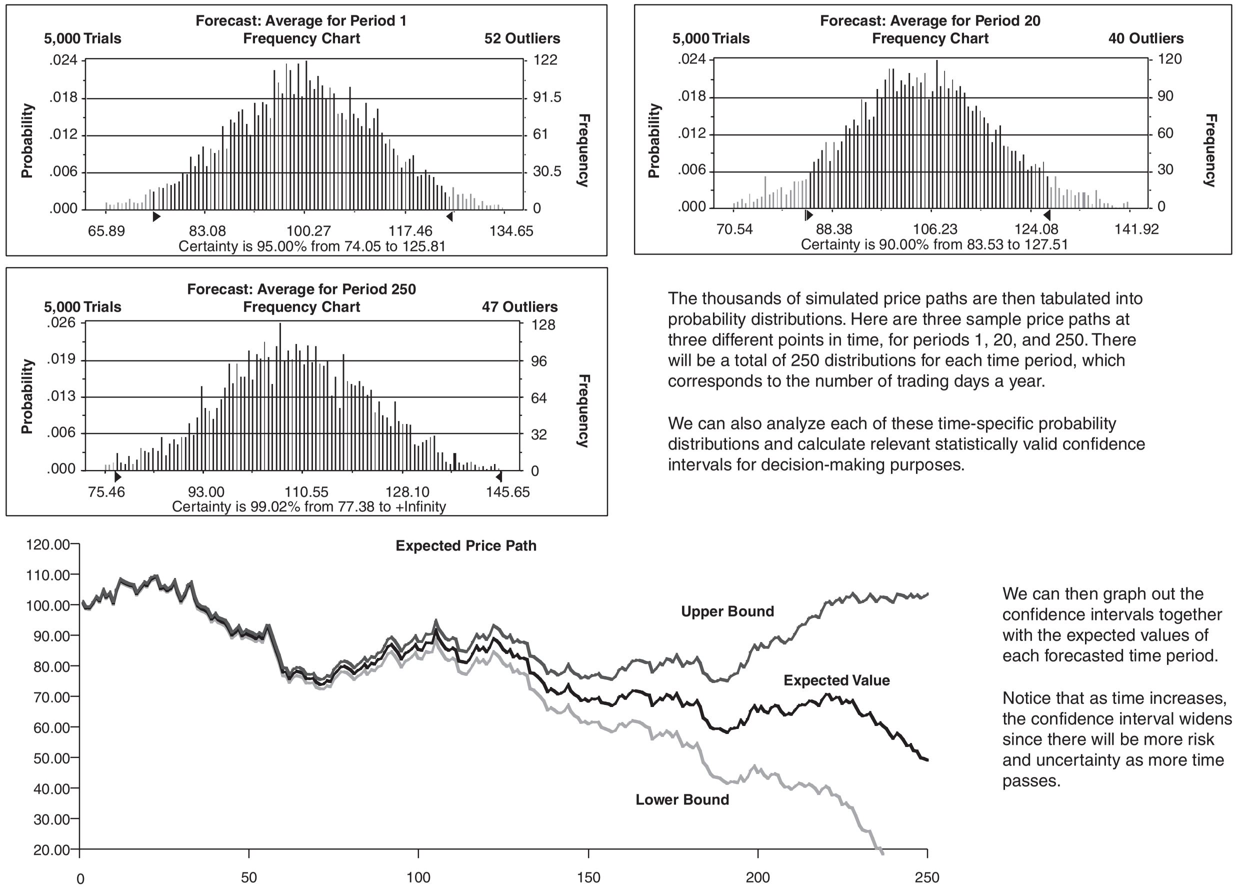


FIGURE 4.4 Conceptualizing the lognormal distribution.

of Monte Carlo simulation for forecasting.[4] The examples in Figure 4.4 show how Monte Carlo simulation can be really complicated, depending on its use. The enclosed CD-ROM's Risk Simulator software has a stochastic process module that applies some of these more complex stochastic forecasting models, including Brownian Motion, mean-reversion, and random-walk models.

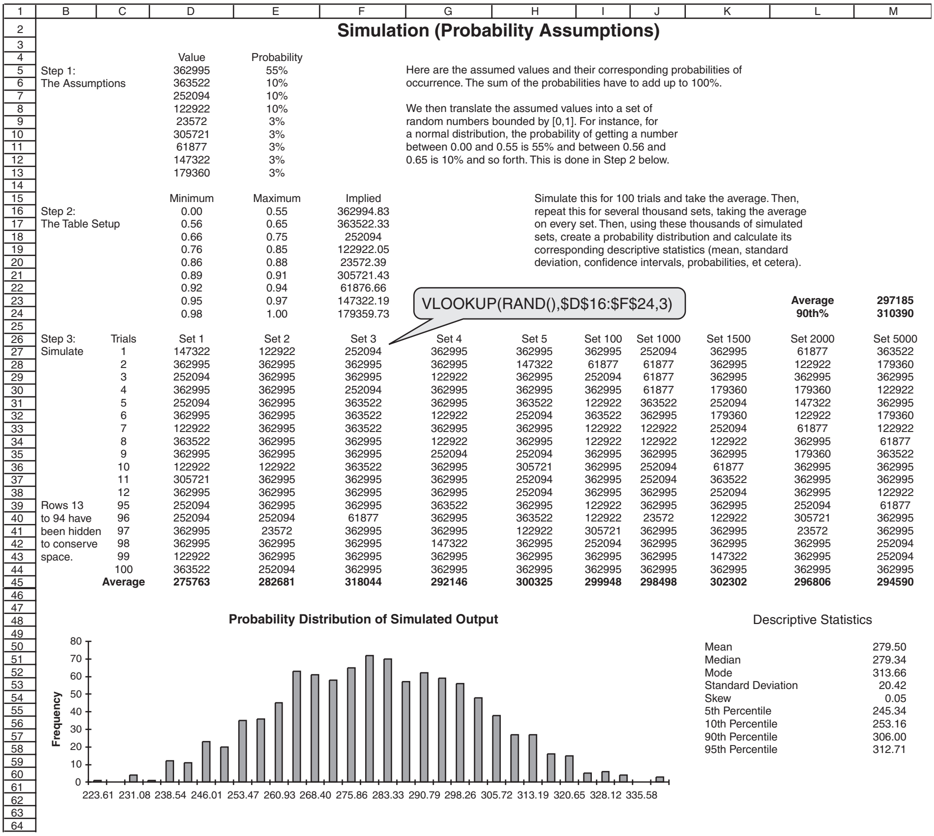
USING RISK SIMULATOR AND EXCEL TO PERFORM SIMULATIONS

Simulations can be performed using Excel. However, more advanced simulation packages such as Risk Simulator perform the task more efficiently and have additional features preset in each simulation. We now present both Monte Carlo parametric simulation and nonparametric bootstrap simulation using Excel and Risk Simulator.

The examples in Figures 4.5 and 4.6 are created using Excel to perform a limited number of simulations on a set of probabilistic assumptions. We assume that having performed a series of scenario analyses, we obtain a set of nine resulting values, complete with their respective probabilities of occurrence. The first step in setting up a simulation in Excel for such a scenario analysis is to understand the function "*RAND( )*" within Excel. This function is simply a random number generator Excel uses to create random numbers from a uniform distribution between 0 and 1. Then translate this 0 to 1 range using the assigned probabilities in our assumption into ranges or bins. For instance, if the value $362,995 occurs with a 55 percent probability, we can create a bin with a range of 0.00 to 0.55. Similarly, we can create a bin range of 0.56 to 0.65 for the next value of $363,522, which occurs 10 percent of the time, and so forth. Based on these ranges and bins, the nonparametric simulation can now be set up.

Figure 4.5 illustrates an example with 5,000 sets of trials. Each set of trials is simulated 100 times; that is, in each simulation trial set, the original numbers are picked randomly with replacement by using the Excel formula *VLOOKUP(RAND(), $D$16:$F$24, 3)*, which picks up the third column of data from the D16 to F24 area by matching the results from the *RAND()* function and data from the first column.

The average of the data sampled is then calculated for each trial set. The distribution of these 5,000 trial sets' averages is obtained and the frequency distribution is shown at the bottom of Figure 4.5. According to the Central Limit Theorem, the average of these sample averages will approach the real true mean of the population at the limit. In addition, the distribution will most likely approach normality when a sufficient set of trials are performed.



1	B	C	D	E	F	G	H	I	J	K	L	M
2					**Simulation (Probability Assumptions)**							
3												
4			Value	Probability								
5	Step 1:		362995	55%		Here are the assumed values and their corresponding probabilities of						
6	The Assumptions		363522	10%		occurrence. The sum of the probabilities have to add up to 100%.						
7			252094	10%								
8			122922	10%		We then translate the assumed values into a set of						
9			23572	3%		random numbers bounded by [0,1]. For instance, for						
10			305721	3%		a normal distribution, the probability of getting a number						
11			61877	3%		between 0.00 and 0.55 is 55% and between 0.56 and						
12			147322	3%		0.65 is 10% and so forth. This is done in Step 2 below.						
13			179360	3%								
14												
15			Minimum	Maximum	Implied		Simulate this for 100 trials and take the average. Then,					
16	Step 2:		0.00	0.55	362994.83		repeat this for several thousand sets, taking the average					
17	The Table Setup		0.56	0.65	363522.33		on every set. Then, using these thousands of simulated					
18			0.66	0.75	252094		sets, create a probability distribution and calculate its					
19			0.76	0.85	122922.05		corresponding descriptive statistics (mean, standard					
20			0.86	0.88	23572.39		deviation, confidence intervals, probabilities, et cetera).					
21			0.89	0.91	305721.43							
22			0.92	0.94	61876.66							
23			0.95	0.97	147322.19	VLOOKUP(RAND(),$D$16:$F$24,3)					**Average**	**297185**
24			0.98	1.00	179359.73						**90th%**	**310390**
25												
26	Step 3:	Trials	Set 1	Set 2	Set 3	Set 4	Set 5	Set 100	Set 1000	Set 1500	Set 2000	Set 5000
27	Simulate	1	147322	122922	252094	362995	362995	362995	252094	362995	61877	363522
28		2	362995	362995	362995	362995	147322	61877	61877	362995	122922	179360
29		3	252094	362995	362995	122922	362995	252094	61877	362995	362995	362995
30		4	362995	362995	252094	362995	362995	362995	61877	179360	179360	122922
31		5	252094	362995	363522	362995	363522	122922	363522	252094	147322	362995
32		6	362995	362995	363522	122922	252094	363522	362995	179360	122922	179360
33		7	122922	362995	363522	362995	362995	122922	122922	252094	61877	122922
34		8	363522	362995	362995	122922	362995	122922	122922	122922	362995	61877
35		9	362995	362995	362995	252094	252094	362995	362995	362995	179360	363522
36		10	122922	122922	363522	362995	305721	362995	252094	61877	362995	362995
37		11	305721	362995	362995	362995	252094	362995	252094	363522	362995	362995
38		12	362995	362995	362995	362995	252094	362995	362995	252094	362995	122922
39	Rows 13	95	252094	362995	362995	363522	362995	122922	362995	362995	252094	61877
40	to 94 have	96	252094	252094	61877	362995	363522	122922	23572	122922	305721	362995
41	been hidden	97	362995	23572	362995	362995	122922	305721	362995	362995	23572	362995
42	to conserve	98	362995	362995	362995	147322	362995	252094	362995	362995	362995	252094
43	space.	99	122922	362995	362995	362995	362995	362995	362995	147322	362995	252094
44		100	363522	252094	362995	362995	362995	362995	362995	362995	362995	362995
45		**Average**	**275763**	**282681**	**318044**	**292146**	**300325**	**299948**	**298498**	**302302**	**296806**	**294590**

Descriptive Statistics

Mean	279.50
Median	279.34
Mode	313.66
Standard Deviation	20.42
Skew	0.05
5th Percentile	245.34
10th Percentile	253.16
90th Percentile	306.00
95th Percentile	312.71

FIGURE 4.5 Simulation using Excel I.

Clearly, running this nonparametric simulation manually in Excel is fairly tedious. An alternative is to use Risk Simulator's custom distribution, which does the same thing but in an infinitely faster and more efficient fashion. Chapter 6, Pandora's Tool Box, illustrates some of these simulation tools in more detail.

Nonparametric simulation is a very powerful tool but it is only applicable if data are available. Clearly, the more data there are, the higher the level of precision and confidence in the simulation results. However, when no data exist or when a valid systematic process underlies the data set (e.g., physics, engineering, economic relationship), parametric simulation may be more appropriate, where exact probabilistic distributions are used.

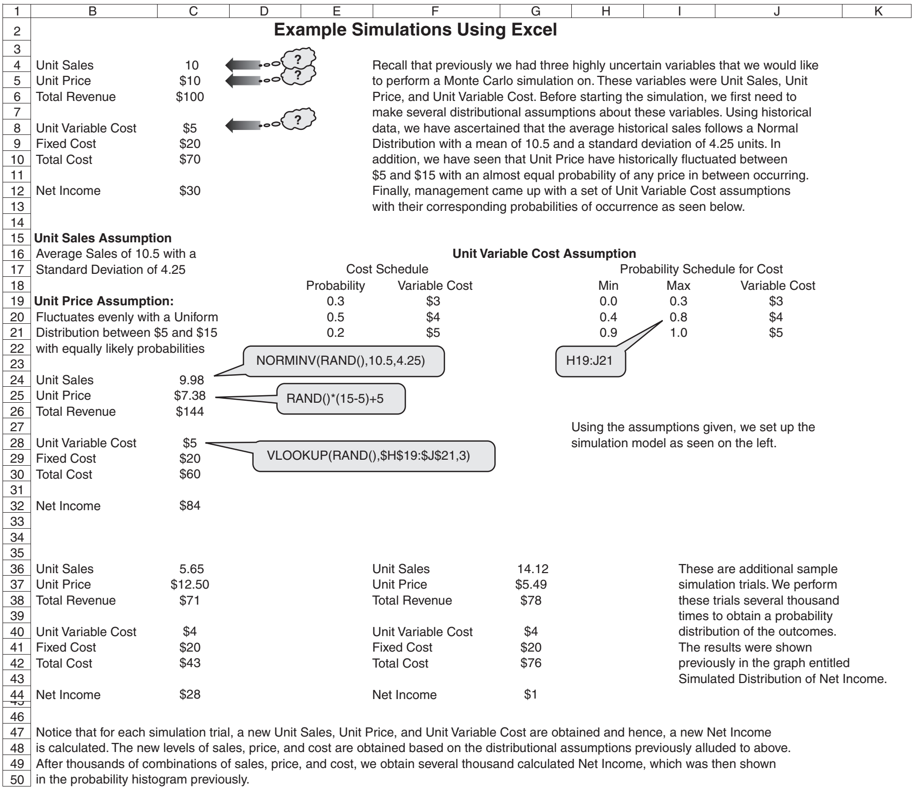


FIGURE 4.6 Simulation using Excel II.

> The *RAND()* function in Excel is used to generate random numbers for a uniform distribution between 0 and 1. *RAND()*(B-A)+A* is used to generate random numbers for a uniform distribution between A and B. *NORMSINV(RAND())* generates random numbers from a standard normal distribution with mean of zero and variance of one.

Using Excel to perform simulations is easy and effective for simple problems. However, when more complicated problems arise, such as the one to be presented next, the use of more specialized simulation packages is warranted. Risk Simulator is such a package. In the example shown in Figure 4.7, the cells for “Revenues,” “Opex,” “FCF/EBITDA Multiple,” and “Revenue Growth Rates” (dark gray) are the assumption cells, where we enter our distributional input assumptions, such as the type of distribution the

Monte Carlo Simulation on Financial Analysis

Project A

		2001	2002	2003	2004	2005
Revenues		$1,010	$1,111	$1,233	$1,384	$1,573
Opex/Revenue Multiple		0.09	0.10	0.11	0.12	0.13
Operating Expenses		$91	$109	$133	$165	$210
EBITDA		$919	$1,002	$1,100	$1,219	$1,363
FCF/EBITDA Multiple		0.20	0.25	0.31	0.40	0.56
Free Cash Flows	($1,200)	$187	$246	$336	$486	$760
Initial Investment	($1,200)					
Revenue Growth Rates		10.00%	11.00%	12.21%	13.70%	15.58%

NPV	$126
IRR	15.68%
Risk Adjusted Discount Rate	12.00%
Growth Rate	3.00%
Terminal Value	$8,692
Terminal Risk Adjustment	30.00%
Discounted Terminal Value	$2,341
Terminal to NPV Ratio	18.52
Payback Period	3.89
Simulated Risk Value	$390

Project B

		2001	2002	2003	2004	2005
Revenues		$1,200	$1,404	$1,683	$2,085	$2,700
Opex/Revenue Multiple		0.09	0.10	0.11	0.12	0.13
Operating Expenses		$108	$138	$181	$249	$361
EBITDA		$1,092	$1,266	$1,502	$1,836	$2,340
FCF/EBITDA Multiple		0.10	0.11	0.12	0.14	0.16
Free Cash Flows	($400)	$109	$139	$183	$252	$364
Initial Investment	($400)					
Revenue Growth Rates		17.00%	19.89%	23.85%	29.53%	38.25%

NPV	$149
IRR	33.74%
Risk Adjusted Discount Rate	19.00%
Growth Rate	3.75%
Terminal Value	$2,480
Terminal Risk Adjustment	30.00%
Discounted Terminal Value	$668
Terminal to NPV Ratio	4.49
Payback Period	2.83
Simulated Risk Value	$122

Project C

		2001	2002	2003	2004	2005
Revenues		$950	$1,069	$1,219	$1,415	$1,678
Opex/Revenue Multiple		0.13	0.15	0.17	0.20	0.24
Operating Expenses		$124	$157	$205	$278	$395
EBITDA		$827	$912	$1,014	$1,136	$1,283
FCF/EBITDA Multiple		0.20	0.25	0.31	0.40	0.56
Free Cash Flows	($1,100)	$168	$224	$309	$453	$715
Initial Investment	($1,100)					
Revenue Growth Rates		12.50%	14.06%	16.04%	18.61%	22.08%

NPV	$29
IRR	15.99%
Risk Adjusted Discount Rate	15.00%
Growth Rate	5.50%
Terminal Value	$7,935
Terminal Risk Adjustment	30.00%
Discounted Terminal Value	$2,137
Terminal to NPV Ratio	74.73
Payback Period	3.88
Simulated Risk Value	$53

Project D

		2001	2002	2003	2004	2005
Revenues		$1,200	$1,328	$1,485	$1,681	$1,932
Opex/Revenue Multiple		0.08	0.08	0.09	0.09	0.10
Operating Expenses		$90	$107	$129	$159	$200
EBITDA		$1,110	$1,221	$1,355	$1,522	$1,732
FCF/EBITDA Multiple		0.14	0.16	0.19	0.23	0.28
Free Cash Flows	($750)	$159	$200	$259	$346	$483
Initial Investment	($750)					
Revenue Growth Rates		10.67%	11.80%	13.20%	14.94%	17.17%

NPV	$26
IRR	21.57%
Risk Adjusted Discount Rate	20.00%
Growth Rate	1.50%
Terminal Value	$2,648
Terminal Risk Adjustment	30.00%
Discounted Terminal Value	$713
Terminal to NPV Ratio	26.98
Payback Period	3.38
Simulated Risk Value	$56

	Implementation Cost	Sharpe Ratio	Weight	Project Cost	Project NPV	Risk Parameter	Payback Period	Technology Level	Tech Mix
Project A	$1,200	0.02	5.14%	$62	$6	29%	3.89	5	0.26
Project B	$400	0.31	25.27%	$101	$38	15%	2.83	3	0.76
Project C	$1,100	0.19	34.59%	$380	$10	21%	3.88	2	0.69
Project D	$750	0.17	35.00%	$263	$9	17%	3.38	4	1.40
Total	$3,450	0.17	100.00%	$806	$63	28%	3.49	3.5	3.11

Constraints:

	Lower Barrier	Upper Barrier	
Budget	$0	$900	(10 percentile at top 900)
Payback Mix	0.10	1.00	
Technology Mix	0.40	4.00	
Per Project Mix	5%	35%	

FIGURE 4.7 Simulation using Risk Simulator.

variable follows and what the parameters are. For instance, we can say that revenues follow a normal distribution with a mean of $1,010 and a standard deviation of $100, based on analyzing historical revenue data for the firm. The net present value (NPV) cells are the forecast output cells, that is, the results of these cells are the results we ultimately wish to analyze. Refer to Chapter 5, Test Driving Risk Simulator, for details on setting up and getting started with using the Risk Simulator software.

QUESTIONS

1. Compare and contrast parametric and nonparametric simulation.
2. What is a stochastic process (e.g., Brownian Motion)?
3. What does the *RAND()* function do in Excel?
4. What does the *NORMSINV()* function do in Excel?
5. What happens when both functions are used together, that is, *NORMSINV(RAND())*?

CHAPTER 5

Test Driving Risk Simulator

This chapter provides the novice risk analyst an introduction to the Risk Simulator software for performing Monte Carlo simulation, a trial version of which is included in the book's CD-ROM. The chapter begins by illustrating what Risk Simulator does and what steps are taken in a Monte Carlo simulation, as well as some of the more basic elements in a simulation analysis. The chapter then continues with how to interpret the results from a simulation and ends with a discussion of correlating variables in a simulation as well as applying precision and error control. As software versions with new enhancements are continually released, please review the software's user manual for more up-to-date details on using the latest version of the software.

The Risk Simulator version 1.1 is a Monte Carlo simulation, forecasting, and optimization software. It is written in Microsoft .NET C# and functions together with Excel as an add-in. This software is also compatible and often used with the Real Options Super Lattice Solver software (see Chapters 12 and 13), both developed by the author. The different functions or modules in both software are:

- The *Simulation Module* allows you to run simulations in your existing Excel-based models, generate and extract simulation forecasts (distributions of results), perform distributional fitting (automatically finding the best-fitting statistical distribution), compute correlations (maintain relationships among simulated random variables), identify sensitivities (creating tornado and sensitivity charts), test statistical hypotheses (finding statistical differences between pairs of forecasts), run bootstrap simulation (testing the robustness of result statistics), and run custom and nonparametric simulations (simulations using historical data without specifying any distributions or their parameters for forecasting without data or applying expert opinion forecasts).
- The *Forecasting Module* can be used to generate automatic time-series forecasts (with and without seasonality and trend), multivariate regressions (modeling relationships among variables), nonlinear extrapolations

(curve fitting), stochastic processes (random walks, mean-reversions, jump-diffusion, and mixed processes), and Box-Jenkins ARIMA (econometric forecasts).

- The *Optimization Module* is used for optimizing multiple decision variables subject to constraints to maximize or minimize an objective, and can be run either as a static optimization, as a dynamic optimization under uncertainty together with Monte Carlo simulation, or as a stochastic optimization. The software can handle linear and nonlinear optimizations with integer and continuous variables.
- The Real Options Super Lattice Solver is another standalone software that complements Risk Simulator, used for solving simple to complex real options problems. See Chapters 12 and 13 for more details on the concept, software, and applications of real options analysis.

GETTING STARTED WITH RISK SIMULATOR

To install the software, insert the accompanying CD-ROM, click on the *Install Risk Simulator* link, and follow the onscreen instructions. You will need to be online to download the latest version of the software. The software requires Windows 2000 or XP, administrative privileges, and Microsoft .Net Framework 1.1 be installed on the computer. Most new computers come with Microsoft .NET Framework 1.1 already preinstalled. However, if an error message pertaining to requiring .NET Framework 1.1 occurs during the installation of Risk Simulator, exit the installation. Then, install the relevant .NET Framework 1.1 software also included in the CD (found in the *DOT NET Framework* folder). Complete the .NET installation, restart the computer, and then reinstall the Risk Simulator software.

Once installation is complete, start Microsoft Excel, and if the installation was successful, you should see an additional *Simulation* item on the menu bar in Excel and a new icon bar as seen in Figure 5.1. Figure 5.2 shows the icon toolbar in more detail. You are now ready to start using the software. The following sections provide step-by-step instructions for using the software. As the software is continually updated and improved, the examples in this book might be slightly different than the latest version downloaded from the Internet.

There is a default 30-day trial license file that comes with the software. To obtain a full corporate license, please contact the author's firm, Real Options Valuation, Inc., at admin@realoptionsvaluation.com. Professors at accredited universities can obtain complimentary renewable semester-long copies of the software both for themselves and for installation in computer labs if both the software and this book are adopted and used in an entire course.

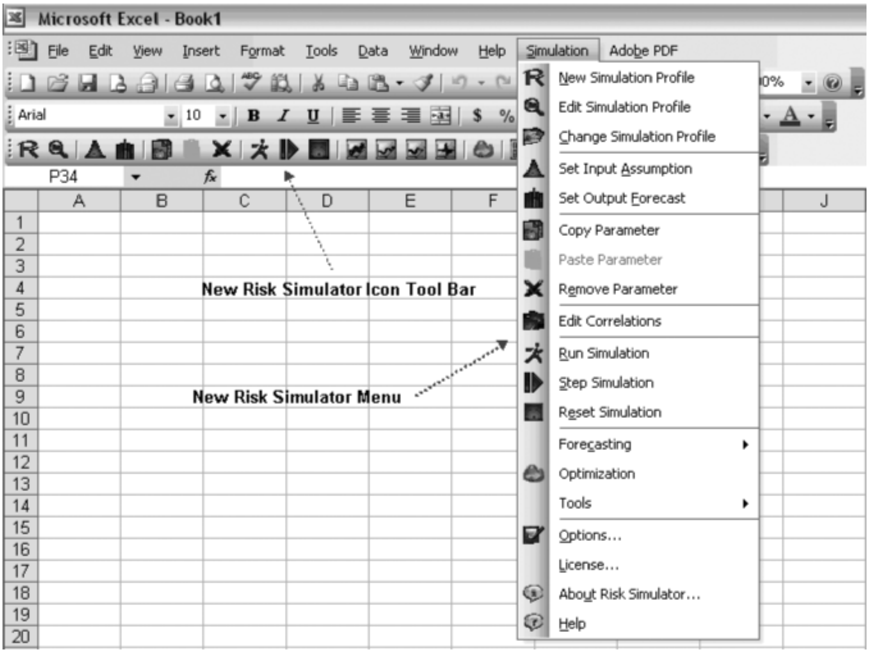


FIGURE 5.1 Risk Simulator menu and icon toolbar.

RUNNING A MONTE CARLO SIMULATION

Typically, to run a simulation in your existing Excel model, the following steps must be performed:

1. Start a new or open an existing simulation profile.
2. Define input assumptions in the relevant cells.
3. Define output forecasts in the relevant cells.
4. Run simulation.
5. Interpret the results.

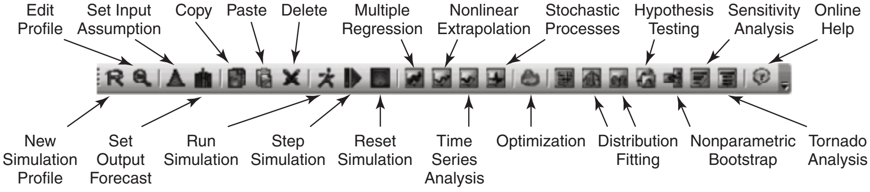


FIGURE 5.2 Risk Simulator icon toolbar.

If desired, and for practice, open the example file called *Basic Simulation Model* and follow along the examples below on creating a simulation. The example file can be found on the start menu at *Start | Real Options Valuation | Risk Simulator | Examples*.

1. Starting a New Simulation Profile

To start a new simulation, you must first create a simulation profile. A simulation profile contains a complete set of instructions on how you would like to run a simulation, that is, all the assumptions, forecasts, simulation run preferences, and so forth. Having profiles facilitates creating multiple scenarios of simulations; that is, using the same exact model, several profiles can be created, each with its own specific simulation assumptions, forecasts, properties, and requirements. The same analyst can create different test scenarios using different distributional assumptions and inputs or multiple users can test their own assumptions and inputs on the same model. Instead of having to make duplicates of the same model, the same model can be used and different simulations can be run through this model *profiling* process.

The following list provides the procedure for starting a new simulation profile:

1. *Start Excel* and create a new or open an existing model (you can use the Basic Simulation Model example to follow along).
2. *Click* on *Simulation | New Simulation Profile*.
3. *Enter* all pertinent information including a title for your simulation (Figure 5.3).

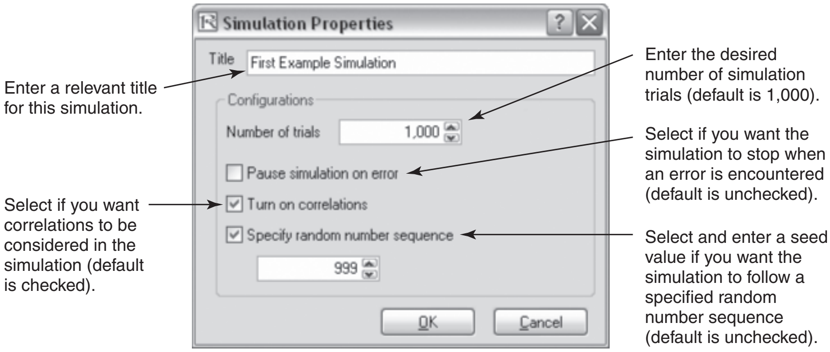


FIGURE 5.3 New simulation profile.

The following are the elements in the new simulation profile dialog box (Figure 5.3):

- *Title.* Specifying a simulation profile name or title allows you to create multiple simulation profiles in a single Excel model, which means that you can now save different simulation scenario profiles within the same model without having to delete existing assumptions and changing them each time a new simulation scenario is required.
- *Number of trials.* The number of simulation trials required is entered; that is, running 1,000 trials means that 1,000 different iterations of outcomes based on the input assumptions will be generated. You can change this number as desired, but the input has to be positive integers. The default number of runs is 1,000 trials.
- *Pause simulation on error.* If checked, the simulation stops every time an error is encountered in the Excel model; that is, if your model encounters a computation error (e.g., some input values generated in a simulation trial may yield a divide-by-zero error in one of your spreadsheet cells), the simulation stops. This feature is important to help audit your model to make sure there are no computational errors in your Excel model. However, if you are sure the model works, then there is no need for this preference to be checked.
- *Turn on correlations.* If checked, correlations between paired input assumptions will be computed. Otherwise, correlations will all be set to zero and a simulation is run assuming no cross-correlations between input assumptions. As an example, applying correlations will yield more accurate results if indeed correlations exist and will tend to yield a lower forecast confidence if negative correlations exist.
- *Specify random number sequence.* By definition simulation yields slightly different results every time it is run by virtue of the random number generation routine in Monte Carlo simulation. This is a theoretical fact in all random number generators. However, when making presentations, sometimes you may require the same results (especially when the report being presented shows one set of results and during a live presentation you would like to show the same results being generated, or when you are sharing models with others and would like the same results to be obtained every time), then check this preference and enter in an initial seed number. The seed number can be any positive integer. Using the same initial seed value, the same number of trials, and the same input assumptions will always yield the same sequence of random numbers, guaranteeing the same final set of results.

Note that once a new simulation profile has been created, you can come back later and modify these selections. In order to do this, make sure that the

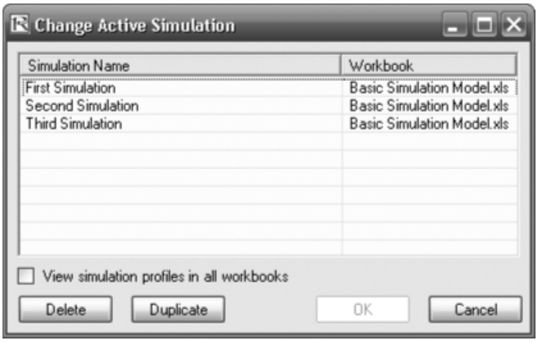


FIGURE 5.4 Change active simulation.

current active profile is the profile you wish to modify; otherwise, click on *Simulation* | *Change Simulation Profile*, select the profile you wish to change and click *OK* (Figure 5.4 shows an example where there are multiple profiles and how to activate, duplicate, or delete a selected profile). Then, click on *Simulation* | *Edit Simulation Profile* and make the required changes.

2. Defining Input Assumptions

The next step is to set input assumptions in your model. Note that assumptions can only be assigned to cells without any equations or functions, that is, typed-in numerical values that are inputs in a model, whereas output forecasts can only be assigned to cells with equations and functions, that is, outputs of a model. Recall that assumptions and forecasts cannot be set unless a simulation profile already exists. Follow this procedure to set new input assumptions in your model:

1. Select the cell you wish to set an assumption on (e.g., cell G8 in the Basic Simulation Model example).
2. Click on *Simulation* | *Set Input Assumption* or click on the set assumption icon in the Risk Simulator icon toolbar.
3. Select the relevant distribution you want, enter the relevant distribution parameters, and hit *OK* to insert the input assumption into your model (Figure 5.5)

Several key areas are worthy of mention in the Assumption Properties. Figure 5.6 shows the different areas:

- *Assumption Name*. This optional area allows you to enter in unique names for the assumptions to help track what each of the assumptions

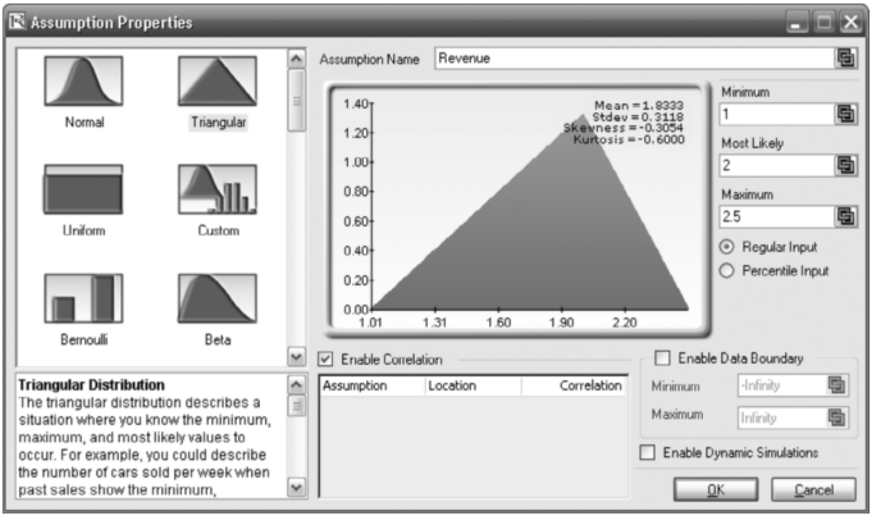


FIGURE 5.5 Setting an input assumption.

represents. Good modeling practice is to use short but precise assumption names.

- *Distribution Gallery*. This area to the left shows all of the different distributions available in the software. To change the views, right click anywhere in the gallery and select large icons, small icons, or list. More than two dozen distributions are available.
- *Input Parameters*. Depending on the distribution selected, the required relevant parameters are shown. You may either enter the parameters directly or link them to specific cells in your worksheet (click on the link icon to link an input parameter to a worksheet cell). Hard coding or typing the parameters is useful when the assumption parameters are assumed not to change. Linking to worksheet cells is useful when the input parameters need to be visible on the worksheets themselves or are allowed to be changed as in a dynamic simulation (where the input parameters themselves are linked to assumptions in the worksheet, creating a multidimensional simulation or simulation of simulations).
- *Data Boundary*. Distributional or data boundaries truncation are typically not used by the average analyst but exist for truncating the distributional assumptions. For instance, if a normal distribution is selected, the theoretical boundaries are between negative infinity and positive infinity. However, in practice, the simulated variable exists only within some smaller range, and this range can then be entered to truncate the distribution appropriately.

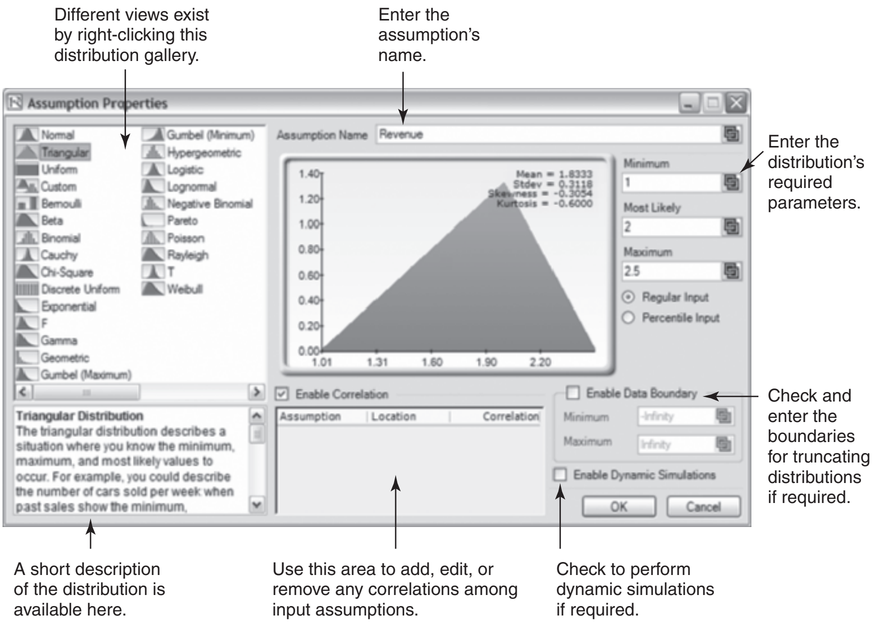


FIGURE 5.6 Assumption properties.

- *Correlations*. Pairwise correlations can be assigned to input assumptions here. If assumptions are required, remember to check the *Turn on Correlations* preference by clicking on *Simulation | Edit Simulation Profile*. See the discussion on correlations later in this chapter for more details about assigning correlations and the effects correlations will have on a model.
- *Short Descriptions*. These exist for each of the distributions in the gallery. The short descriptions explain when a certain distribution is used as well as the input parameter requirements. See the section in the appendix, Understanding Probability Distributions, for details about each distribution type available in the software.

Note: If you are following along with the example, continue by setting another assumption on cell G9. This time use the Uniform distribution with a minimum value of 0.9 and a maximum value of 1.1. Then, proceed to defining the output forecasts in the next step.

3. Defining Output Forecasts

The next step is to define output forecasts in the model. Forecasts can only be defined on output cells with equations or functions.

Use the following procedure to define the forecasts:

1. Select the cell on which you wish to set an assumption (e.g., cell G10 in the Basic Simulation Model example).
2. Click on *Simulation | Set Output Forecast* or click on the set forecast icon on the Risk Simulator icon toolbar.
3. Enter the relevant information and click *OK*.

Figure 5.7 illustrates the set forecast properties:

- *Forecast Name*. Specify the name of the forecast cell. This is important because when you have a large model with multiple forecast cells, naming the forecast cells individually allows you to access the right results quickly. Do not underestimate the importance of this simple step. Good modeling practice is to use short but precise assumption names.
- *Forecast Precision*. Instead of relying on a guesstimate of how many trials to run in your simulation, you can set up precision and error controls. When an error–precision combination has been achieved in the simulation, the simulation will pause and inform you of the precision achieved, making the number of simulation trials an automated process and not making you rely on guesses of the required number of trials to simulate. Review the section on error and precision control for more specific details.
- *Show Forecast Window*. This property allows the user to show or not show a particular forecast window. The default is to always show a forecast chart.

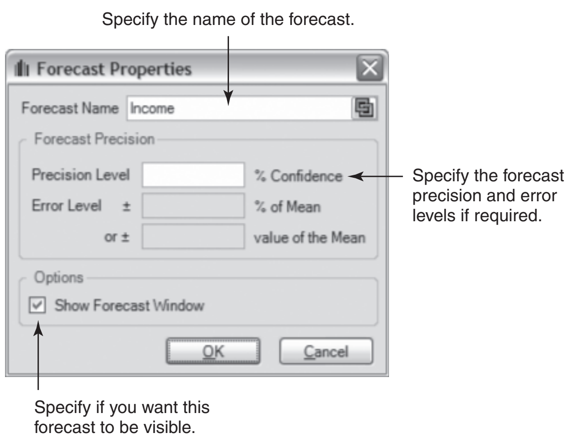


FIGURE 5.7 Set output forecast.

4. Run Simulation

If everything looks right, simply click on *Simulation | Run Simulation* or click on the *Run* icon on the Risk Simulator toolbar, and the simulation will proceed. You may also reset a simulation after it has run to rerun it (*Simulation | Reset Simulation* or the reset icon on the toolbar), or to pause it during a run. Also, the *step* function (*Simulation | Step Simulation* or the step icon on the toolbar) allows you to simulate a single trial, one at a time, useful for educating others on simulation (i.e., you can show that at each trial, all the values in the assumption cells are being replaced and the entire model is recalculated each time).

5. Interpreting the Forecast Results

The final step in Monte Carlo simulation is to interpret the resulting forecast charts. Figures 5.8 to 5.15 show the forecast chart and the corresponding statistics generated after running the simulation. Typically, the following sections on the forecast window are important in interpreting the results of a simulation:

- *Forecast Chart.* The forecast chart shown in Figure 5.8 is a probability histogram that shows the frequency counts of values occurring and the total number of trials simulated. The vertical bars show the frequency of a particular x value occurring out of the total number of trials, while the cumulative frequency (smooth line) shows the total probabilities of all values at and below x occurring in the forecast.
- *Forecast Statistics.* The forecast statistics shown in Figure 5.9 summarize the distribution of the forecast values in terms of the four moments

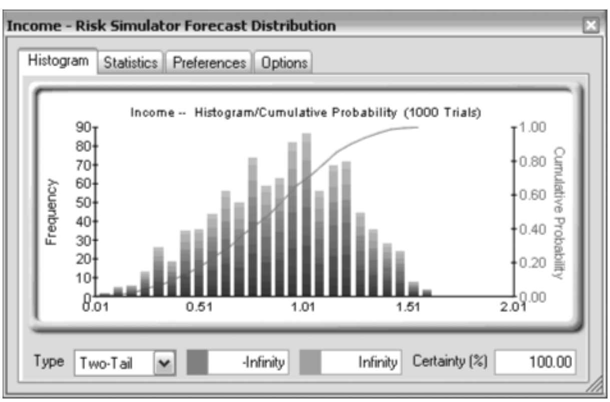


FIGURE 5.8 Forecast chart.

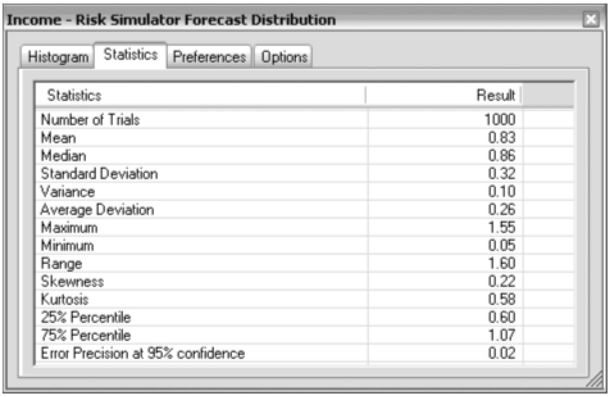


FIGURE 5.9 Forecast statistics.

of a distribution. See The Statistics of Risk in Chapter 2 for more details on what some of these statistics mean. You can rotate between the histogram and statistics tab by depressing the space bar.

- *Preferences.* The preferences tab in the forecast chart (Figure 5.10) allows you to change the look and feel of the charts. For instance, if *Always Show Window On Top* is selected, the forecast charts will always be visible regardless of what other software is running on your computer. The *Semitransparent When Inactive* is a powerful option used to compare or overlay multiple forecast charts at once (e.g., enable this option on several forecast charts and drag them on top of one another to visually see the similarities or differences. *Histogram Resolution* allows

FIGURE 5.10 Forecast chart preferences.

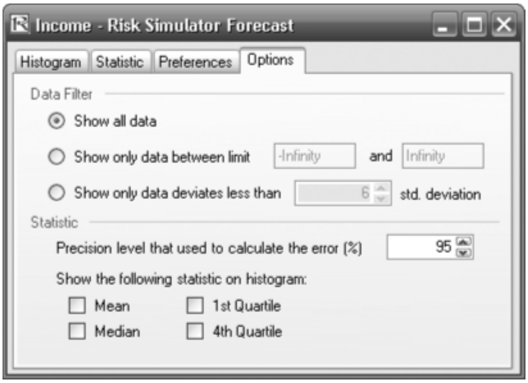


FIGURE 5.11 Forecast chart options.

you to change the number of bins of the histogram, anywhere from 5 bins to 100 bins. Also, the *Data Update Interval* section allows you to control how fast the simulation runs versus how often the forecast chart is updated; that is, if you wish to see the forecast chart updated at almost every trial, this feature will slow down the simulation as more memory is being allocated to updating the chart versus running the simulation. This section is merely a user preference and in no way changes the results of the simulation, just the speed of completing the simulation. You can also click on *Close All* and *Minimize All* to close or minimize the existing forecast windows.

- *Options*. This forecast chart option (Figure 5.11) allows you to show all the forecast data or to filter in or out values that fall within some specified interval, or within some standard deviation that you choose. Also, the precision level can be set here for this specific forecast to show the error levels in the statistics view. See the section Correlations and Precision Control for more details.

USING FORECAST CHARTS AND CONFIDENCE INTERVALS

In forecast charts, you can determine the probability of occurrence called *confidence intervals*; that is, given two values, what are the chances that the outcome will fall between these two values? Figure 5.12 illustrates that there is a 90 percent probability that the final outcome (in this case, the level of income) will be between \$0.2647 and \$1.3230. The two-tailed confidence interval can be obtained by first selecting *Two-Tail* as the type, entering the desired certainty value (e.g., 90), and hitting *Tab* on the keyboard. The two

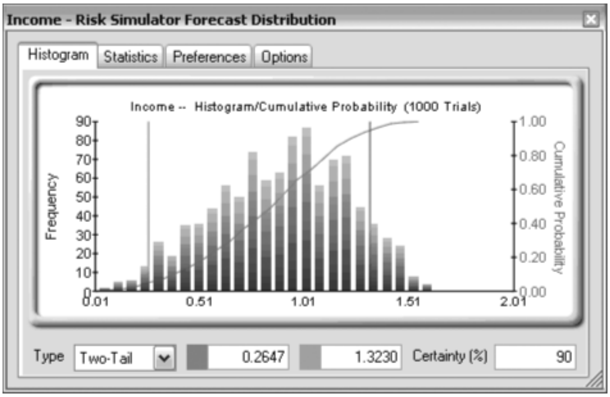


FIGURE 5.12 Forecast chart two-tailed confidence interval.

computed values corresponding to the certainty value will then be displayed. In this example, there is a 5 percent probability that income will be below $0.2647 and another 5 percent probability that income will be above $1.3230; that is, the two-tailed confidence interval is a symmetrical interval centered on the median or 50th percentile value. Thus, both tails will have the same probability.

Alternatively, a one-tail probability can be computed. Figure 5.13 shows a *Left-Tail* selection at 95 percent confidence (i.e., choose *Left-Tail* as the type, enter 95 as the certainty level, and hit *Tab* on the keyboard). This means that there is a 95 percent probability that the income will be below $1.3230 (i.e., 95 percent on the left-tail of $1.3230) or a 5 percent probability that

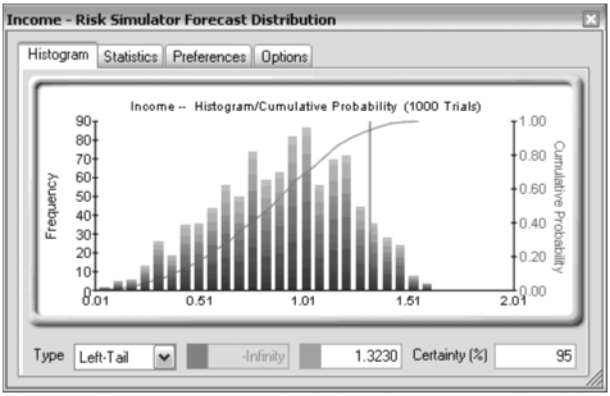


FIGURE 5.13 Forecast chart one-tailed confidence interval.

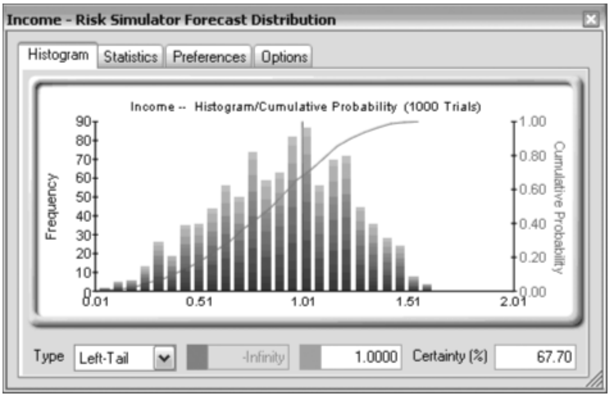


FIGURE 5.14 Forecast chart left tail probability evaluation.

income will be above \$1.3230, corresponding perfectly with the results seen in Figure 5.12.

In addition to evaluating the confidence interval (i.e., given a probability level and finding the relevant income values), you can determine the probability of a given income value (Figure 5.14). For instance, what is the probability that income will be less than \$1? To do this, select the *Left-Tail* probability type, enter 1 into the value input box, and hit *Tab*. The corresponding certainty will then be computed (in this case, there is a 67.70 percent probability income will be below \$1).

For the sake of completeness, you can select the *Right-Tail* probability type and enter the value 1 in the value input box, and hit *Tab* (Figure 5.15).

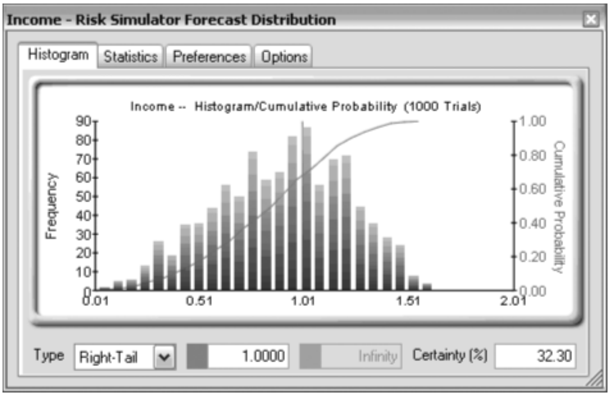


FIGURE 5.15 Forecast chart right tail probability evaluation.

The resulting probability indicates the right-tail probability past the value 1, that is, the probability of income exceeding \$1 (in this case, we see that there is a 32.30 percent probability of income exceeding \$1).

Note that the forecast window is resizable by clicking on and dragging the bottom right corner of the forecast window. Finally, it is always advisable that before rerunning a simulation, the current simulation should be reset by selecting *Simulation | Reset Simulation*.

CORRELATIONS AND PRECISION CONTROL

The Basics of Correlations

The correlation coefficient is a measure of the strength and direction of the relationship between two variables, and can take on any values between –1.0 and +1.0; that is, the correlation coefficient can be decomposed into its direction or sign (positive or negative relationship between two variables) and the magnitude or strength of the relationship (the higher the absolute value of the correlation coefficient, the stronger the relationship).

The correlation coefficient can be computed in several ways. The first approach is to manually compute the correlation coefficient r of a pair of variables x and y using:

$$r_{x,y} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - \left(\sum x_i\right)^2}\sqrt{n\sum y_i^2 - \left(\sum y_i\right)^2}}$$

The second approach is to use Excel's *CORREL* function. For instance, if the 10 data points for x and y are listed in cells A1:B10, then the Excel function to use is *CORREL (A1:A10, B1:B10)*. The third approach is to run Risk Simulator's *Multi-Variable Distributional Fitting Tool* and the resulting correlation matrix will be computed and displayed.

It is important to note that correlation does not imply causation. Two completely unrelated random variables might display some correlation, but this does not imply any causation between the two (e.g., sunspot activity and events in the stock market are correlated, but there is no causation between the two).

There are two general types of correlations: parametric and nonparametric correlations. Pearson's correlation coefficient is the most common correlation measure, and is usually referred to simply as the correlation coefficient. However, Pearson's correlation is a parametric measure, which means that it requires both correlated variables to have an underlying normal distribution and that the relationship between the variables is linear. When these conditions are violated, which is often the case in Monte Carlo simulation, the nonparametric counterparts become more important. Spearman's rank correlation and Kendall's tau are the two nonparametric alternatives. The Spearman correlation is most commonly used and is most appropriate when applied in the context of Monte Carlo simulation—there is no dependence on normal distributions or linearity, meaning that correlations between different variables with different distribution can be applied. In order to compute the Spearman correlation, first rank all the x and y variable values and then apply the Pearson's correlation computation.

In the case of Risk Simulator, the correlation used is the more robust nonparametric Spearman's rank correlation. However, to simplify the simulation process and to be consistent with Excel's correlation function, the correlation user inputs required are the Pearson's correlation coefficient. Risk Simulator will then apply its own algorithms to convert them into Spearman's rank correlation, thereby simplifying the process.

Applying Correlations in Risk Simulator

Correlations can be applied in Risk Simulator in several ways:

- When defining assumptions, simply enter the correlations into the correlation grid in the Distribution Gallery.
- With existing data, run the Multi-Variable Distribution Fitting tool to perform distributional fitting and to obtain the correlation matrix between pairwise variables. If a simulation profile exists, the assumptions fitted will automatically contain the relevant correlation values.
- With the use of a direct-input correlation matrix, click on *Simulation | Edit Correlations* to view and edit the correlation matrix used in the simulation.

Note that the correlation matrix must be positive definite; that is, the correlation must be mathematically valid. For instance, suppose you are trying to correlate three variables: grades of graduate students in a particular year, the number of beers they consume a week, and the number of hours they study a week. One would assume that the following correlation relationships exist:

Grades and Beer: –	The more they drink, the lower the grades (no show on exams).
Grades and Study: +	The more they study, the higher the grades.
Beer and Study: –	The more they drink, the less they study (drunk and partying all the time).

However, if you input a negative correlation between Grades and Study and assuming that the correlation coefficients have high magnitudes, the correlation matrix will be nonpositive definite. It would defy logic, correlation requirements, and matrix mathematics. However, smaller coefficients can sometimes still work even with the bad logic. When a nonpositive definite or bad correlation matrix is entered, Risk Simulator automatically informs you of the error and offers to adjust these correlations to something that is semipositive definite while still maintaining the overall structure of the correlation relationship (the same signs as well as the same relative strengths).

The Effects of Correlations in Monte Carlo Simulation

Although the computations required to correlate variables in a simulation is complex, the resulting effects are fairly clear. Table 5.1 shows a simple correlation model (Correlation Effects Model in the example folder). The calculation for revenue is simply price multiplied by quantity. The same model is replicated for no correlations, positive correlation (+0.9), and negative correlation (–0.9) between price and quantity.

The resulting statistics are shown in Figure 5.16. Notice that the standard deviation of the model without correlations is 0.23, compared to 0.30 for the positive correlation, and 0.12 for the negative correlation; that is, for simple models with positive relationships (e.g., additions and multiplications), negative correlations tend to reduce the average spread of the distribution and create a tighter and more concentrated forecast distribution as compared to positive correlations with larger average spreads. However, the mean remains relatively stable. This implies that correlations do little to change the expected value of projects but can reduce or increase a project's risk. Recall in financial theory that negatively correlated variables, projects, or assets when combined in a portfolio tend to create a diversification effect where the overall risk is reduced. Therefore, we see a smaller standard deviation for the negatively correlated model.

Table 5.2 illustrates the results after running a simulation, extracting the raw data of the assumptions, and computing the correlations between the

TABLE 5.1 Simple Correlation Model

	Without Correlation	Positive Correlation	Negative Correlation
Price	$2.00	$2.00	$2.00
Quantity	1.00	1.00	1.00
Revenue	$2.00	$2.00	$2.00

Revenue No Correlation - Risk Simulator Forecast Distribution

Histogram | Statistics | Preferences | Options

Statistics	Result
Number of Trials	1000
Mean	2.00
Median	2.01
Standard Deviation	0.23
Variance	0.05
Average Deviation	0.18
Maximum	2.78
Minimum	1.17
Range	1.61
Skewness	0.02
Kurtosis	0.12
25% Percentile	1.84
75% Percentile	2.16
Error Precision at 95% confidence	0.01

Revenue Positive Correlation - Risk Simulator Forecast Distribution

Histogram | Statistics | Preferences | Options

Statistics	Result
Number of Trials	1000
Mean	2.00
Median	1.99
Standard Deviation	0.30
Variance	0.09
Average Deviation	0.25
Maximum	2.87
Minimum	1.33
Range	1.54
Skewness	0.20
Kurtosis	0.60
25% Percentile	1.76
75% Percentile	2.23
Error Precision at 95% confidence	0.01

Revenue Negative Correlation - Risk Simulator Forecast Distribution

Histogram | Statistics | Preferences | Options

Statistics	Result
Number of Trials	1000
Mean	1.99
Median	2.00
Standard Deviation	0.12
Variance	0.02
Average Deviation	0.10
Maximum	2.47
Minimum	1.40
Range	1.07
Skewness	0.57
Kurtosis	1.34
25% Percentile	1.91
75% Percentile	2.07
Error Precision at 95% confidence	0.00

FIGURE 5.16 Correlation results.

TABLE 5.2 Correlations Recovered

Price Negative Correlation	Quantity Negative Correlation	Correlation	Price Positive Correlation	Quantity Positive Correlation	Correlation
676	145	–0.90	102	158	0.89
368	452		461	515	
264	880		515	477	
235	877		874	833	
122	711		769	792	
490	641		481	471	
336	638		627	446	
495	383		82	190	
241	568		659	674	
651	571		188	286	
854	59		458	439	
66	950		981	972	
707	262		528	569	
943	186		865	812	

variables. The table shows that the input assumptions are recovered in the simulation; that is, you enter +0.9 and –0.9 correlations and the resulting simulated values have the same correlations. Clearly there will be minor differences from one simulation run to another, but when enough trials are run, the resulting recovered correlations approach those that were input.

Precision and Error Control

One very powerful tool in Monte Carlo simulation is that of precision control. For instance, how many trials are considered sufficient to run in a complex model? Precision control takes the guesswork out of estimating the relevant number of trials by allowing the simulation to stop if the level of the prespecified precision is reached.

The precision control functionality lets you set how precise you want your forecast to be. Generally speaking, as more trials are calculated, the confidence interval narrows and the statistics become more accurate. The precision control feature in Risk Simulator uses the characteristic of confidence intervals to determine when a specified accuracy of a statistic has been reached. For each forecast, you can specify the specific confidence interval for the precision level.

Make sure that you do not confuse three very different terms: error, precision, and confidence. Although they sound similar, the concepts are significantly different from one another. A simple illustration is in order. Suppose you are a taco shell manufacturer and are interested in finding out how many broken taco shells there are on average in a single box of 100 shells. One way to do this is to collect a sample of prepackaged boxes of 100 taco shells, open them, and count how many of them are actually broken. You manufacture 1 million boxes a day (this is your *population*), but you randomly open only 10 boxes (this is your *sample* size, also known as your number of *trials* in a simulation). The number of broken shells in each box is as follows: 24, 22, 4, 15, 33, 32, 4, 1, 45, and 2. The calculated average number of broken shells is 18.2. Based on these 10 samples or trials, the average is 18.2 units, while based on the sample, the 80 percent confidence interval is between 2 and 33 units (that is, 80 percent of the time, the number of broken shells is between 2 and 33 *based on this sample size or number of trials run*). However, how sure are you that 18.2 is the correct average? Are 10 trials sufficient to establish this average and confidence level? The confidence interval between 2 and 33 is too wide and too variable. Suppose you require a more accurate average value where the error is ±2 taco shells 90 percent of the time—this means that if you open *all* 1 million boxes manufactured in a day, 900,000 of these boxes will have broken taco shells on average at some mean unit ±2 tacos. How many more taco shell boxes would you then need to sample (or trials run) to obtain this level of precision? Here, the 2 tacos is the error level while the 90 percent is the level of precision. If sufficient numbers of trials are run, then the 90 percent confidence interval will be identical to the 90 percent precision level, where a more precise measure of the average is obtained such that 90 percent of the time, the error, and hence, the confidence will be ±2 tacos. As an example, say the average is 20 units, then the 90 percent confidence interval will be between 18 and 22 units, where this interval is precise 90 percent of the time, where in opening all 1 million boxes, 900,000 of them will have between 18 and 22 broken tacos. Stated differently, we have a 10 percent error level with respect to the mean (i.e., 2 divided by 20) at the 90 percent confidence level. The terms *percent error* and *percent confidence* level are standard terms used in statistics and in Risk Simulator.

The number of trials required to hit this precision is based on the sampling error equation of

$$\bar{x} \pm Z \frac{s}{\sqrt{n}}$$

where $Z \frac{s}{\sqrt{n}}$ is the error of 2 tacos

- $\bar{x}$ is the sample average
- Z is the standard-normal Z-score obtained from the 90 percent precision level
- s is the sample standard deviation
- n is the number of trials required to hit this level of error with the specified precision.[1]

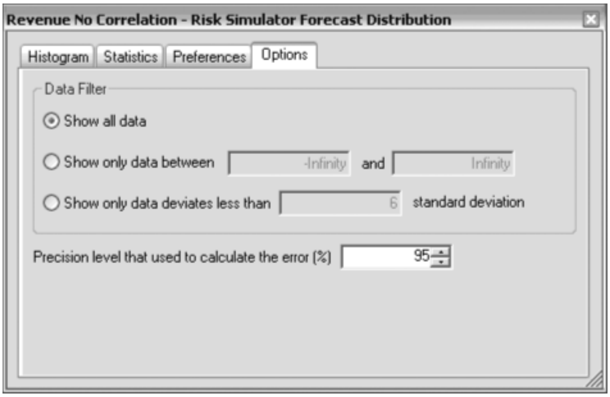


FIGURE 5.17 Setting the forecast's precision level.

Figures 5.17 and 5.18 illustrate how precision control can be performed on multiple simulated forecasts in Risk Simulator. This feature prevents the user from having to decide how many trials to run in a simulation and eliminates all possibilities of guesswork. Figure 5.18 shows that there is a 0.01 percent error with respect to the mean at a 95 percent confidence level.

Using the simple techniques outlined in this chapter, you are well on your way to running Monte Carlo simulations with Risk Simulator. Later chapters continue with additional techniques and tools available in Risk Simulator to further enhance your analysis.

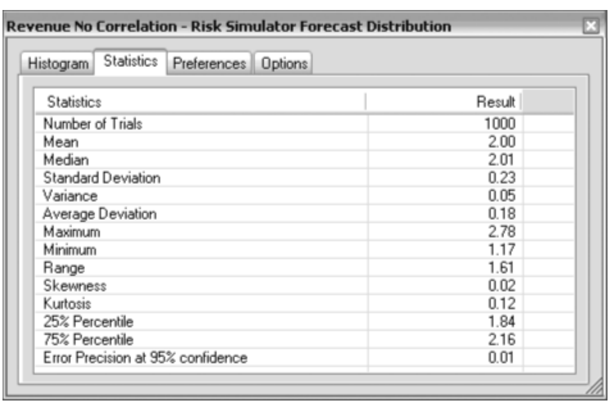


FIGURE 5.18 Computing the error.

APPENDIX—UNDERSTANDING PROBABILITY DISTRIBUTIONS

This chapter demonstrates the power of Monte Carlo simulation, but in order to get started with simulation, one first needs to understand the concept of probability distributions. This appendix continues with the use of the author's Risk Simulator software and shows how simulation can be very easily and effortlessly implemented in an existing Excel model. A limited trial version of the Risk Simulator software is available in the enclosed CD-ROM (to obtain a permanent version, please visit the author's web site at www.realoptionsvaluation.com). Professors can obtain free semester-long computer lab licenses for their students and themselves if this book and the simulation/options valuation software are used and taught in an entire class.

To begin to understand probability, consider this example: You want to look at the distribution of nonexempt wages within one department of a large company. First, you gather raw data—in this case, the wages of each nonexempt employee in the department. Second, you organize the data into a meaningful format and plot the data as a frequency distribution on a chart. To create a frequency distribution, you divide the wages into group intervals and list these intervals on the chart's horizontal axis. Then you list the number or frequency of employees in each interval on the chart's vertical axis. Now you can easily see the distribution of nonexempt wages within the department.

A glance at the chart illustrated in Figure 5.19 reveals that the employees earn from \$7.00 to \$9.00 per hour. You can chart this data as a probability distribution. A probability distribution shows the number of employees

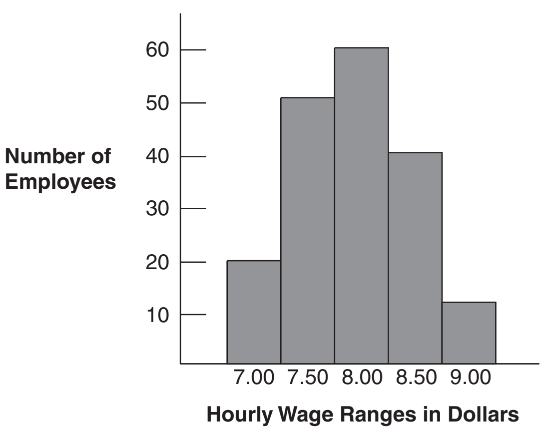


FIGURE 5.19 Frequency histogram I.

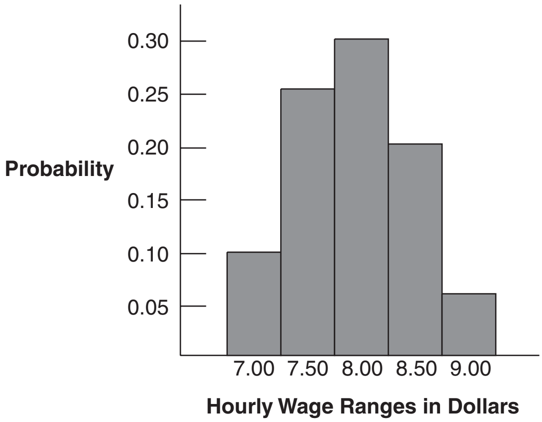


FIGURE 5.20 Frequency histogram II.

in each interval as a fraction of the total number of employees. To create a probability distribution, you divide the number of employees in each interval by the total number of employees and list the results on the chart's vertical axis.

The chart in Figure 5.20 shows the number of employees in each wage group as a fraction of all employees; you can estimate the likelihood or probability that an employee drawn at random from the whole group earns a wage within a given interval. For example, assuming the same conditions exist at the time the sample was taken, the probability is 0.20 (a one in five chance) that an employee drawn at random from the whole group earns $8.50 an hour.

Probability distributions are either discrete or continuous. *Discrete probability distributions* describe distinct values, usually integers, with no intermediate values and are shown as a series of vertical bars. A discrete distribution, for example, might describe the number of heads in four flips of a coin as 0, 1, 2, 3, or 4. *Continuous probability distributions* are actually mathematical abstractions because they assume the existence of every possible intermediate value between two numbers; that is, a continuous distribution assumes there is an infinite number of values between any two points in the distribution. However, in many situations, you can effectively use a continuous distribution to approximate a discrete distribution even though the continuous model does not necessarily describe the situation exactly.

Selecting a Probability Distribution

Plotting data is one method for selecting a probability distribution. The following steps provide another process for selecting probability distributions that best describe the uncertain variables in your spreadsheets.

To select the correct probability distribution, use the following steps:

1. Look at the variable in question. List everything you know about the conditions surrounding this variable. You might be able to gather valuable information about the uncertain variable from historical data. If historical data are not available, use your own judgment, based on experience, listing everything you know about the uncertain variable.
2. Review the descriptions of the probability distributions.
3. Select the distribution that characterizes this variable. A distribution characterizes a variable when the conditions of the distribution match those of the variable.

Alternatively, if you have historical, comparable, contemporaneous, or forecast data, you can use Risk Simulator's distributional fitting modules to find the best statistical fit for your existing data. This fitting process will apply some advanced statistical techniques to find the best distribution and its relevant parameters that describe the data.

Probability Density Functions, Cumulative Distribution Functions, and Probability Mass Functions

In mathematics and Monte Carlo simulation, a probability density function (PDF) represents a *continuous* probability distribution in terms of integrals. If a probability distribution has a density of $f(x)$, then intuitively the infinitesimal interval of $[x, x + dx]$ has a probability of $f(x)\,dx$. The PDF therefore can be seen as a smoothed version of a probability histogram; that is, by providing an empirically large sample of a continuous random variable repeatedly, the histogram using very narrow ranges will resemble the random variable's PDF. The probability of the interval between $[a, b]$ is given by

$$\int_a^b f(x)dx$$

which means that the total integral of the function f must be 1.0. *It is a common mistake to think of f(a) as the probability of a.* This is incorrect. In fact, $f(a)$ can sometimes be larger than 1—consider a uniform distribution between 0.0 and 0.5. The random variable x within this distribution will have $f(x)$ greater than 1. The probability in reality is the function $f(x)dx$ discussed previously, where dx is an infinitesimal amount.

The cumulative distribution function (CDF) is denoted as $F(x) = P(X \le x)$, indicating the probability of X taking on a less than or equal value to x.

Every CDF is monotonically increasing, is continuous from the right, and at the limits, has the following properties:

$$\lim_{x \to -\infty} F(x) = 0 \quad \text{and} \quad \lim_{x \to +\infty} F(x) = 1$$

Further, the CDF is related to the PDF by

$$F(b) - F(a) = P(a \leq X \leq b) = \int_a^b f(x)dx$$

where the PDF function f is the derivative of the CDF function F.

In probability theory, a probability mass function or PMF gives the probability that a *discrete* random variable is exactly equal to some value. The PMF differs from the PDF in that the values of the latter, defined only for continuous random variables, are not probabilities; rather, its integral over a set of possible values of the random variable is a probability. A random variable is discrete if its probability distribution is discrete and can be characterized by a PMF. Therefore, X is a discrete random variable if

$$\sum_u P(X = u) = 1$$

as u runs through all possible values of the random variable X.

Discrete Distributions

Following is a detailed listing of the different types of probability distributions that can be used in Monte Carlo simulation. This listing is included in the appendix for the reader's reference.

Bernoulli or Yes/No Distribution The Bernoulli distribution is a discrete distribution with two outcomes (e.g., head or tails, success or failure, 0 or 1). The Bernoulli distribution is the binomial distribution with one trial and can be used to simulate Yes/No or Success/Failure conditions. This distribution is the fundamental building block of other more complex distributions. For instance:

- *Binomial distribution.* Bernoulli distribution with higher number of n total trials and computes the probability of x successes within this total number of trials.
- *Geometric distribution.* Bernoulli distribution with higher number of trials and computes the number of failures required before the first success occurs.
- *Negative binomial distribution.* Bernoulli distribution with higher number of trials and computes the number of failures before the xth success occurs.

The mathematical constructs for the Bernoulli distribution are as follows:

$$P(x) = \begin{cases} 1-p & \text{for } x = 0 \\ p & \text{for } x = 1 \end{cases}$$

or

$$P(x) = p^x(1-p)^{1-x}$$

$$Mean = p$$

$$Standard\ Deviation = \sqrt{p(1-p)}$$

$$Skewness = \frac{1-2p}{\sqrt{p(1-p)}}$$

$$Excess\ Kurtosis = \frac{6p^2-6p+1}{p(1-p)}$$

The probability of success (p) is the only distributional parameter. Also, it is important to note that there is only one trial in the Bernoulli distribution, and the resulting simulated value is either 0 or 1.

Input requirements:

Probability of success > 0 and < 1 (that is, $0.0001 \le p \le 0.9999$)

Binomial Distribution The binomial distribution describes the number of times a particular event occurs in a fixed number of trials, such as the number of heads in 10 flips of a coin or the number of defective items out of 50 items chosen.

The three conditions underlying the binomial distribution are:

1. For each trial, only two outcomes are possible that are mutually exclusive.
2. The trials are independent—what happens in the first trial does not affect the next trial.
3. The probability of an event occurring remains the same from trial to trial.

$$P(x) = \frac{n!}{x!(n-x)!}p^x(1-p)^{(n-x)} \text{ for } n > 0;\ x = 0, 1, 2, \ldots \text{n; and } 0 < p < 1$$

$$Mean = np$$

$$Standard\ Deviation = \sqrt{np(1-p)}$$

$$Skewness = \frac{1-2p}{\sqrt{np(1-p)}}$$

$$Excess\ Kurtosis = \frac{6p^2-6p+1}{np(1-p)}$$

The probability of success (p) and the integer number of total trials (n) are the distributional parameters. The number of successful trials is denoted x. It is important to note that probability of success (p) of 0 or 1 are trivial conditions and do not require any simulations, and, hence, are not allowed in the software.

Input requirements:

Probability of success > 0 and < 1 (that is, $0.0001 \le p \le 0.9999$).

Number of trials ≥ 1 or positive integers and ≤ 1,000 (for larger trials, use the normal distribution with the relevant computed binomial mean and standard deviation as the normal distribution's parameters).

Discrete Uniform The discrete uniform distribution is also known as the *equally likely outcomes* distribution, where the distribution has a set of N elements, and each element has the same probability. This distribution is related to the uniform distribution, but its elements are discrete and not continuous.

The mathematical constructs for the discrete uniform distribution are as follows:

$$P(x) = \frac{1}{N} \text{ ranked value}$$

$$Mean = \frac{N+1}{2} \text{ ranked value}$$

$$Standard\ Deviation = \sqrt{\frac{(N-1)(N+1)}{12}} \text{ ranked value}$$

$$Skewness = 0 \text{ (that is, the distribution is perfectly symmetrical)}$$

$$Excess\ Kurtosis = \frac{-6(N^2+1)}{5(N-1)(N+1)} \text{ ranked value}$$

Input requirements:

Minimum < Maximum and both must be integers (negative integers and zero are allowed)

Geometric Distribution The geometric distribution describes the number of trials until the first successful occurrence, such as the number of times you need to spin a roulette wheel before you win.

The three conditions underlying the geometric distribution are:

1. The number of trials is not fixed.
2. The trials continue until the first success.
3. The probability of success is the same from trial to trial.

The mathematical constructs for the geometric distribution are as follows:

$$P(x) = p(1-p)^{x-1} \text{ for } 0 < p < 1 \text{ and } x = 1, 2, \ldots, n$$

$$Mean = \frac{1}{p} - 1$$

$$Standard\ Deviation = \sqrt{\frac{1-p}{p^2}}$$

$$Skewness = \frac{2-p}{\sqrt{1-p}}$$

$$Excess\ Kurtosis = \frac{p^2 - 6p + 6}{1-p}$$

The probability of success (p) is the only distributional parameter. The number of successful trials simulated is denoted x, which can only take on positive integers.

Input requirements:

Probability of success > 0 and < 1 (that is, $0.0001 \leq p \leq 0.9999$). It is important to note that probability of success (p) of 0 or 1 are trivial conditions and do not require any simulations, and, hence, are not allowed in the software.

Hypergeometric Distribution The hypergeometric distribution is similar to the binomial distribution in that both describe the number of times a particular event occurs in a fixed number of trials. The difference is that binomial distribution trials are independent, whereas hypergeometric distribution trials change the probability for each subsequent trial and are called *trials without replacement*. For example, suppose a box of manufactured parts is known to contain some defective parts. You choose a part from the box, find it is defective, and remove the part from the box. If you choose another part from the box, the probability that it is defective is somewhat lower than for the first part because you have removed a defective part. If you had replaced the defective part, the probabilities would have remained the same, and the process would have satisfied the conditions for a binomial distribution.

The three conditions underlying the hypergeometric distribution are:

1. The total number of items or elements (the population size) is a fixed number, a finite population. The population size must be less than or equal to 1,750.
2. The sample size (the number of trials) represents a portion of the population.
3. The known initial probability of success in the population changes after each trial.

The mathematical constructs for the hypergeometric distribution are as follows:

$$P(x) = \frac{\dfrac{(N_x)!}{x!(N_x - x)!}\dfrac{(N - N_x)!}{(n-x)!(N - N_x - n + x)!}}{\dfrac{N!}{n!(N-n)!}} \text{ for } x = Max(n - (N - N_x), 0), \ldots, Min(n, N_x)$$

$$Mean = \frac{N_x n}{N}$$

$$Standard\ Deviation = \sqrt{\frac{(N - N_x)N_x n(N-n)}{N^2(N-1)}}$$

$$Skewness = \frac{(N - 2N_x)(N - 2n)}{N-2}\sqrt{\frac{N-1}{(N-N_x)N_x n(N-n)}}$$

$$Excess\ Kurtosis = \frac{V(N, N_x, n)}{(N - N_x)\ N_x n(-3+N)(-2+N)(-N+n)} \text{ where}$$

$$\begin{aligned}V(N, Nx, n) = {} & (N - N_x)^3 - (N - N_x)^5 + 3(N - N_x)^2 N_x - 6(N - N_x)^3 N_x \\ & +(N - N_x)^4 Nx + 3(N - N_x)\ N_x^2 - 12(N - N_x)^2 N_x^2 + 8(N - N_x)^3 N_x^2 + N_x^3 \\ & -6(N - N_x)\ N_x^3 + 8(N - N_x)^2 N_x^3 + (N - N_x)\ N_x^4 - N_x^5 - 6(N - N_x)^3 N_x \\ & +6(N - N_x)^4 N_x + 18(N - N_x)^2 N_x n - 6(N - N_x)^3 N_x n + 18(N - N_x)\ N_x^2 n \\ & -24(N - N_x)^2 N_x^2 n - 6(N - N_x)^3 n - 6(N - N_x)\ N_x^3 n + 6N_x^4 n + 6(N - N_x)^2 n^2 \\ & -6(N - N_x)^3 n^2 - 24(N - N_x)\ N_x n^2 + 12(N - N_x)^2 N_x n^2 + 6N_x^2 n^2 \\ & +12(N - N_x)\ N_x^2 n^2 - 6N_x^3 n^2\end{aligned}$$

The number of items in the population (N), trials sampled (n), and number of items in the population that have the successful trait (N_x) are the distributional parameters. The number of successful trials is denoted x.

Input requirements:

Population $\geq$ 2 and integer
Trials > 0 and integer

Successes > 0 and integer
Population > Successes
Trials < Population
Population < 1,750

Negative Binomial Distribution The negative binomial distribution is useful for modeling the distribution of the number of trials until the *r*th successful occurrence, such as the number of sales calls you need to make to close a total of 10 orders. It is essentially a *superdistribution* of the geometric distribution. This distribution shows the probabilities of each number of trials in excess of *r* to produce the required success *r*.

The three conditions underlying the negative binomial distribution are:

1. The number of trials is not fixed.
2. The trials continue until the *r*th success.
3. The probability of success is the same from trial to trial.

The mathematical constructs for the negative binomial distribution are as follows:

$$P(x) = \frac{(x+r-1)!}{(r-1)!\,x!}p^r(1-p)^x \text{ for } x = r, r+1, \ldots; \text{ and } 0 < p < 1$$

$$Mean = \frac{r(1-p)}{p}$$

$$Standard\ Deviation = \sqrt{\frac{r(1-p)}{p^2}}$$

$$Skewness = \frac{2-p}{\sqrt{r(1-p)}}$$

$$Excess\ Kurtosis = \frac{p^2 - 6p + 6}{r(1-p)}$$

The probability of success (p) and required successes (r) are the distributional parameters.

Input requirements:

Successes required must be positive integers > 0 and < 8,000.

Probability of success > 0 and < 1 (that is, $0.0001 \le p \le 0.9999$). It is important to note that probability of success (p) of 0 or 1 are trivial conditions and do not require any simulations, and, hence, are not allowed in the software.

Poisson Distribution The Poisson distribution describes the number of times an event occurs in a given interval, such as the number of telephone calls per minute or the number of errors per page in a document.

The three conditions underlying the Poisson distribution are:

1. The number of possible occurrences in any interval is unlimited.
2. The occurrences are independent. The number of occurrences in one interval does not affect the number of occurrences in other intervals.
3. The average number of occurrences must remain the same from interval to interval.

The mathematical constructs for the Poisson are as follows:

$$P(x) = \frac{e^{-\lambda}\lambda^{x}}{x!} \text{ for } x \text{ and } \lambda > 0$$

$$Mean = \lambda$$

$$Standard\ Deviation = \sqrt{\lambda}$$

$$Skewness = \frac{1}{\sqrt{\lambda}}$$

$$Excess\ Kurtosis = \frac{1}{\lambda}$$

Rate (λ) is the only distributional parameter.

Input requirements:

Rate > 0 and ≤ 1,000 (that is, 0.0001 ≤ rate ≤ 1,000)

Continuous Distributions

Beta Distribution The beta distribution is very flexible and is commonly used to represent variability over a fixed range. One of the more important applications of the beta distribution is its use as a conjugate distribution for the parameter of a Bernoulli distribution. In this application, the beta distribution is used to represent the uncertainty in the probability of occurrence of an event. It is also used to describe empirical data and predict the random behavior of percentages and fractions, as the range of outcomes is typically between 0 and 1.

The value of the beta distribution lies in the wide variety of shapes it can assume when you vary the two parameters, alpha and beta. If the parameters are equal, the distribution is symmetrical. If either parameter is 1 and the other parameter is greater than 1, the distribution is J-shaped. If alpha is less than beta, the distribution is said to be positively skewed (most of the values are near the minimum value). If alpha is greater than beta, the distribution is negatively skewed (most of the values are near the maximum value).

The mathematical constructs for the beta distribution are as follows:

$$f(x) = \frac{(x)^{(\alpha-1)}(1-x)^{(\beta-1)}}{\left[\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}\right]} \text{ for } \alpha > 0;\ \beta > 0;\ x > 0$$

$$Mean = \frac{\alpha}{\alpha+\beta}$$

$$Standard\ Deviation = \sqrt{\frac{\alpha\beta}{(\alpha+\beta)^2(1+\alpha+\beta)}}$$

$$Skewness = \frac{2(\beta-\alpha)\sqrt{1+\alpha+\beta}}{(2+\alpha+\beta)\sqrt{\alpha\beta}}$$

$$Excess\ Kurtosis = \frac{3(\alpha+\beta+1)[\alpha\,\beta(\alpha+\beta-6)+2(\alpha+\beta)^2]}{\alpha\,\beta(\alpha+\beta+2)(\alpha+\beta+3)} - 3$$

Alpha (α) and beta (β) are the two distributional shape parameters, and Γ is the gamma function.

The two conditions underlying the beta distribution are:

1. The uncertain variable is a random value between 0 and a positive value.
2. The shape of the distribution can be specified using two positive values.

Input requirements:

Alpha and beta > 0 and can be any positive value

Cauchy Distribution or Lorentzian Distribution or Breit–Wigner Distribution The Cauchy distribution, also called the Lorentzian distribution or Breit–Wigner distribution, is a continuous distribution describing resonance behavior. It also describes the distribution of horizontal distances at which a line segment tilted at a random angle cuts the x-axis.

The mathematical constructs for the Cauchy or Lorentzian distribution are as follows:

$$f(x) = \frac{1}{\pi}\frac{\gamma / 2}{(x-m)^2 + \gamma^2 / 4}$$

The Cauchy distribution is a special case where it does not have any theoretical moments (mean, standard deviation, skewness, and kurtosis) as they are all undefined.

Mode location (m) and scale (γ) are the only two parameters in this distribution. The location parameter specifies the peak or mode of the distribution, while the scale parameter specifies the half-width at half-maximum of the distribution. In addition, the mean and variance of a Cauchy or Lorentzian distribution are undefined.

In addition, the Cauchy distribution is the Student's t distribution with only 1 degree of freedom. This distribution is also constructed by taking the ratio of two standard normal distributions (normal distributions with a mean of zero and a variance of one) that are independent of one another.

Input requirements:

Location can be any value
Scale > 0 and can be any positive value

Chi-Square Distribution The chi-square distribution is a probability distribution used predominantly in hypothesis testing, and is related to the gamma distribution and the standard normal distribution. For instance, the sums of independent normal distributions are distributed as a chi-square (χ^2) with k degrees of freedom:

$$Z_1^2 + Z_2^2 + \ldots + Z_k^2 \overset{d}{\sim} \chi_k^2$$

The mathematical constructs for the chi-square distribution are as follows:

$$f(x) = \frac{2^{(-k/2)}}{\Gamma(k\,/\,2)} x^{k/2-1} e^{-x/2} \text{ for all } x > 0$$

$$Mean = k$$

$$Standard\ Deviation = \sqrt{2k}$$

$$Skewness = 2\sqrt{\frac{2}{k}}$$

$$Excess\ Kurtosis = \frac{12}{k}$$

The gamma function is written as Γ. Degrees of freedom k is the only distributional parameter.

The chi-square distribution can also be modeled using a gamma distribution by setting the shape parameter as $k/2$ and scale as $2S^2$ where S is the scale.

Input requirements:

Degrees of freedom > 1 and must be an integer < 1,000

Exponential Distribution The exponential distribution is widely used to describe events recurring at random points in time, such as the time between failures of electronic equipment or the time between arrivals at a service booth. It is related to the Poisson distribution, which describes the number of occurrences of an event in a given interval of time. An important characteristic of the exponential distribution is the "memoryless" property, which means that the future lifetime of a given object has the same distribution, regardless of the time it existed. In other words, time has no effect on future outcomes.

The mathematical constructs for the exponential distribution are as follows:

$$f(x) = \lambda e^{-\lambda x} \text{ for } x \geq 0; \lambda > 0$$

$$Mean = \frac{1}{\lambda}$$

$$Standard\ Deviation = \frac{1}{\lambda}$$

$Skewness = 2$ (this value applies to all success rate λ inputs)

$Excess\ Kurtosis = 6$ (this value applies to all success rate λ inputs)

Success rate (λ) is the only distributional parameter. The number of successful trials is denoted x.

The condition underlying the exponential distribution is:

1. The exponential distribution describes the amount of time between occurrences.

   Input requirements:

   Rate > 0 and ≤ 300

Extreme Value Distribution or Gumbel Distribution The extreme value distribution (Type 1) is commonly used to describe the largest value of a response over a period of time, for example, in flood flows, rainfall, and earthquakes. Other applications include the breaking strengths of materials, construction design, and aircraft loads and tolerances. The extreme value distribution is also known as the Gumbel distribution.

The mathematical constructs for the extreme value distribution are as follows:

$$f(x) = \frac{1}{\beta} z e^{-Z} \text{ where } z = e^{\frac{x-m}{\beta}} \text{ for } \beta > 0; \text{ and any value of } x \text{ and } m$$

$$Mean = m + 0.577215\beta$$

$$Standard\ Deviation = \sqrt{\frac{1}{6}\pi^2\beta^2}$$

$$Skewness = \frac{12\sqrt{6}(1.2020569)}{\pi^3} = 1.13955 \text{ (this applies for all values of mode and scale)}$$

$$Excess\ Kurtosis = 5.4 \text{ (this applies for all values of mode and scale)}$$

Mode (m) and scale (β) are the distributional parameters.

There are two standard parameters for the extreme value distribution: mode and scale. The mode parameter is the most likely value for the variable (the highest point on the probability distribution). The scale parameter is a number greater than 0. The larger the scale parameter, the greater the variance.

Input requirements:

Mode can be any value
Scale > 0

F Distribution or Fisher–Snedecor Distribution The F distribution, also known as the Fisher–Snedecor distribution, is another continuous distribution used most frequently for hypothesis testing. Specifically, it is used to test the statistical difference between two variances in analysis of variance tests and likelihood ratio tests. The F distribution with the numerator degree of freedom n and denominator degree of freedom m is related to the chi-square distribution in that:

$$\frac{\chi_n^2 \,/\, n}{\chi_m^2 \,/\, m} \overset{d}{\sim} F_{n,m} \text{ or } f(x) = \frac{\Gamma\left(\frac{n+m}{2}\right)\left(\frac{n}{m}\right)^{n/2} x^{n/2-1}}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{m}{2}\right)\left[x\left(\frac{n}{m}\right)+1\right]^{(n+m)/2}}$$

$$Mean = \frac{m}{m-2}$$

$$Standard\ Deviation = \frac{2m^2(m+n-2)}{n(m-2)^2(m-4)} \text{ for all } m > 4$$

$$Skewness = \frac{2(m+2n-2)}{m-6}\sqrt{\frac{2(m-4)}{n(m+n-2)}}$$

$$Excess\ Kurtosis = \frac{12(-16+20m-8m^2+m^3+44n-32mn+5m^2n-22n^2+5mn^2}{n(m-6)(m-8)(n+m-2)}$$

The numerator degree of freedom n and denominator degree of freedom m are the only distributional parameters.

Input requirements:

Degrees of freedom numerator and degrees of freedom denominator both > 0 integers

Gamma Distribution (Erlang Distribution) The gamma distribution applies to a wide range of physical quantities and is related to other distributions: lognormal, exponential, Pascal, Erlang, Poisson, and chi-square. It is used in meteorological processes to represent pollutant concentrations and precipitation quantities. The gamma distribution is also used to measure the time between the occurrence of events when the event process is not completely random. Other applications of the gamma distribution include inventory control, economic theory, and insurance risk theory.

The gamma distribution is most often used as the distribution of the amount of time until the rth occurrence of an event in a Poisson process. When used in this fashion, the three conditions underlying the gamma distribution are:

1. The number of possible occurrences in any unit of measurement is not limited to a fixed number.
2. The occurrences are independent. The number of occurrences in one unit of measurement does not affect the number of occurrences in other units.
3. The average number of occurrences must remain the same from unit to unit.

The mathematical constructs for the gamma distribution are as follows:

$$f(x) = \frac{\left(\frac{x}{\beta}\right)^{\alpha-1} e^{-\frac{x}{\beta}}}{\Gamma(\alpha)\beta} \text{ with any value of } \alpha > 0 \text{ and } \beta > 0$$

$$Mean = \alpha\beta$$

$$Standard\ Deviation = \sqrt{\alpha\beta^2}$$

$$Skewness = \frac{2}{\sqrt{\alpha}}$$

$$Excess\ Kurtosis = \frac{6}{\alpha}$$

Shape parameter alpha (α) and scale parameter beta (β) are the distributional parameters, and Γ is the gamma function.

When the alpha parameter is a positive integer, the gamma distribution is called the Erlang distribution, used to predict waiting times in queuing systems, where the Erlang distribution is the sum of independent and identically distributed random variables each having a memoryless exponential distribution. Setting n as the number of these random variables, the mathematical construct of the Erlang distribution is:

$$f(x) = \frac{x^{n-1}e^{-x}}{(n-1)!} \text{ for all } x > 0 \text{ and all positive integers of } n$$

Input requirements:

Scale beta > 0 and can be any positive value
Shape alpha ≥ 0.05 and any positive value
Location can be any value

Logistic Distribution The logistic distribution is commonly used to describe growth, that is, the size of a population expressed as a function of a time variable. It also can be used to describe chemical reactions and the course of growth for a population or individual.

The mathematical constructs for the logistic distribution are as follows:

$$f(x) = \frac{e^{\frac{\mu - x}{\alpha}}}{\alpha\left[1 + e^{\frac{\mu - x}{\alpha}}\right]^2} \text{ for any value of } \alpha \text{ and } \mu$$

$$Mean = \mu$$

$$Standard\ Deviation = \sqrt{\frac{1}{3}\pi^2\alpha^2}$$

$Skewness = 0$ (this applies to all mean and scale inputs)

$Excess\ Kurtosis = 1.2$ (this applies to all mean and scale inputs)

Mean (μ) and scale (α) are the distributional parameters.

There are two standard parameters for the logistic distribution: mean and scale. The mean parameter is the average value, which for this distribution is the same as the mode, because this distribution is symmetrical. The scale parameter is a number greater than 0. The larger the scale parameter, the greater the variance.

Input requirements:

Scale > 0 and can be any positive value
Mean can be any value

Lognormal Distribution The lognormal distribution is widely used in situations where values are positively skewed, for example, in financial analysis for security valuation or in real estate for property valuation, and where values cannot fall below zero.

Stock prices are usually positively skewed rather than normally (symmetrically) distributed. Stock prices exhibit this trend because they cannot fall below the lower limit of zero but might increase to any price without limit. Similarly, real estate prices illustrate positive skewness and are lognormally distributed as property values cannot become negative.

The three conditions underlying the lognormal distribution are:

1. The uncertain variable can increase without limits but cannot fall below zero.
2. The uncertain variable is positively skewed, with most of the values near the lower limit.
3. The natural logarithm of the uncertain variable yields a normal distribution.

Generally, if the coefficient of variability is greater than 30 percent, use a lognormal distribution. Otherwise, use the normal distribution.

The mathematical constructs for the lognormal distribution are as follows:

$$f(x) = \frac{1}{x\sqrt{2\pi}\,\ln(\sigma)} e^{-\frac{[\ln(x)-\ln(\mu)]^2}{2[\ln(\sigma)]^2}} \quad \text{for } x > 0; \mu > 0 \text{ and } \sigma > 0$$

$$Mean = \exp\left(\mu + \frac{\sigma^2}{2}\right)$$

$$Standard\ Deviation = \sqrt{\exp\left(\sigma^2 + 2\mu\right)\left[\exp\left(\sigma^2\right) - 1\right]}$$

$$Skewness = \left[\sqrt{\exp\left(\sigma^2\right) - 1}\right](2 + \exp(\sigma^2))$$

$$Excess\ Kurtosis = \exp\left(4\sigma^2\right) + 2\exp\left(3\sigma^2\right) + 3\exp\left(2\sigma^2\right) - 6$$

Mean (μ) and standard deviation (σ) are the distributional parameters.
Input requirements:

Mean and standard deviation both > 0 and can be any positive value

Lognormal Parameter Sets By default, the lognormal distribution uses the arithmetic mean and standard deviation. For applications for which historical data are available, it is more appropriate to use either the logarithmic mean and standard deviation, or the geometric mean and standard deviation.

Normal Distribution The normal distribution is the most important distribution in probability theory because it describes many natural phenomena, such as people's IQs or heights. Decision makers can use the normal distribution to describe uncertain variables such as the inflation rate or the future price of gasoline.

The three conditions underlying the normal distribution are:

1. Some value of the uncertain variable is the most likely (the mean of the distribution).
2. The uncertain variable could as likely be above the mean as it could be below the mean (symmetrical about the mean).
3. The uncertain variable is more likely to be in the vicinity of the mean than further away.

The mathematical constructs for the normal distribution are as follows:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \text{ for all values of } x \text{ and } \mu; \text{ while } \sigma > 0$$

$$Mean = \mu$$

$$Standard\ Deviation = \sigma$$

$Skewness = 0$ (this applies to all inputs of mean and standard deviation)

$Excess\ Kurtosis = 0$ (this applies to all inputs of mean and standard deviation)

Mean (μ) and standard deviation (σ) are the distributional parameters. Input requirements:

Standard deviation > 0 and can be any positive value
Mean can be any value

Pareto Distribution The Pareto distribution is widely used for the investigation of distributions associated with such empirical phenomena as city population sizes, the occurrence of natural resources, the size of companies, personal incomes, stock price fluctuations, and error clustering in communication circuits.

The mathematical constructs for the pareto are as follows:

$$f(x) = \frac{\beta L^{\beta}}{x^{(1+\beta)}} \text{ for } x > L$$

$$Mean = \frac{\beta L}{\beta - 1}$$

$$Standard\ Deviation = \sqrt{\frac{\beta L^2}{(\beta - 1)^2 (\beta - 2)}}$$

$$Skewness = \sqrt{\frac{\beta - 2}{\beta}} \left[\frac{2(\beta + 1)}{\beta - 3} \right]$$

$$Excess\ Kurtosis = \frac{6(\beta^3 + \beta^2 - 6\beta - 2)}{\beta(\beta - 3)(\beta - 4)}$$

Location (L) and shape (β) are the distributional parameters.

There are two standard parameters for the pareto distribution: location and shape. The location parameter is the lower bound for the variable. After you select the location parameter, you can estimate the shape parameter. The shape parameter is a number greater than 0, usually greater than 1. The larger the shape parameter, the smaller the variance and the thicker the right tail of the distribution.

Input requirements:

Location > 0 and can be any positive value
Shape ≥ 0.05

Student's t Distribution The Student's t distribution is the most widely used distribution in hypothesis testing. This distribution is used to estimate the mean of a normally distributed population when the sample size is small, and is used to test the statistical significance of the difference between two sample means or confidence intervals for small sample sizes.

The mathematical constructs for the t distribution are as follows:

$$f(t) = \frac{\Gamma[(r+1)/2]}{\sqrt{r\pi}\Gamma[r/2]}(1 + t^2 / r)^{-(r+1)/2}$$

where $t = \frac{x - \bar{x}}{s}$ and Γ is the gamma function

$Mean = 0$ (this applies to all degrees of freedom r except if the distribution is shifted to another nonzero central location)

$$Standard\ Deviation = \sqrt{\frac{r}{r - 2}}$$

$$Skewness = 0 \text{ (this applies to all degrees of freedom } r\text{)}$$

$$Excess\ Kurtosis = \frac{6}{r-4} \text{ for all } r > 4$$

Degree of freedom r is the only distributional parameter.

The t distribution is related to the F-distribution as follows: The square of a value of t with r degrees of freedom is distributed as F with 1 and r degrees of freedom. The overall shape of the probability density function of the t distribution also resembles the bell shape of a normally distributed variable with mean 0 and variance 1, except that it is a bit lower and wider or is leptokurtic (fat tails at the ends and peaked center). As the number of degrees of freedom grows (say, above 30), the t distribution approaches the normal distribution with mean 0 and variance 1.

Input requirements:

Degrees of freedom ≥ 1 and must be an integer

Triangular Distribution The triangular distribution describes a situation where you know the minimum, maximum, and most likely values to occur. For example, you could describe the number of cars sold per week when past sales show the minimum, maximum, and usual number of cars sold.

The three conditions underlying the triangular distribution are:

1. The minimum number of items is fixed.
2. The maximum number of items is fixed.
3. The most likely number of items falls between the minimum and maximum values, forming a triangular-shaped distribution, which shows that values near the minimum and maximum are less likely to occur than those near the most likely value.

The mathematical constructs for the triangular distribution are as follows:

$$f(x) = \begin{cases} \dfrac{2(x - Min)}{(Max - Min)(Likely - Min)} & \text{for } Min < x < Likely \\ \dfrac{2(Max - x)}{(Max - Min)(Max - Likely)} & \text{for } Likely < x < Max \end{cases}$$

$$Mean = \frac{1}{3}(Min + Likely + Max)$$

$$Standard\ Deviation = \sqrt{\frac{1}{18}(Min^2 + Likely^2 + Max^2 - MinMax - MinLikely - MaxLikely)}$$

$$Skewness = \frac{\sqrt{2}(Min + Max - 2Likely)(2Min - Max - Likely)(Min - 2Max + Likely)}{5(Min^2 + Max^2 + Likely^2 - MinMax - MinLikely - MaxLikely)^{3/2}}$$

$$Excess\ Kurtosis = -0.6 \text{ (this applies to all inputs of } Min, Max, \text{and } Likely\text{)}$$

Minimum value (*Min*), most likely value (*Likely*), and maximum value (*Max*) are the distributional parameters.

Input requirements:

$Min \leq Most\ Likely \leq Max$ and can also take any value
However, $Min < Max$ and can also take any value

Uniform Distribution With the uniform distribution, all values fall between the minimum and maximum and occur with equal likelihood.

The three conditions underlying the uniform distribution are:

1. The minimum value is fixed.
2. The maximum value is fixed.
3. All values between the minimum and maximum occur with equal likelihood.

The mathematical constructs for the uniform distribution are as follows:

$$f(x) = \frac{1}{Max - Min} \text{ for all values such that } Min < Max$$

$$Mean = \frac{Min + Max}{2}$$

$$Standard\ Deviation = \sqrt{\frac{(Max - Min)^2}{12}}$$

$Skewness = 0$ (this applies to all inputs of *Min* and *Max*)

$Excess\ Kurtosis = -1.2$ (this applies to all inputs of *Min* and *Max*)

Maximum value (*Max*) and minimum value (*Min*) are the distributional parameters.

Input requirements:

$Min < Max$ and can take any value

Weibull Distribution (Rayleigh Distribution) The Weibull distribution describes data resulting from life and fatigue tests. It is commonly used to describe failure time in reliability studies as well as the breaking strengths of materials in reliability and quality control tests. Weibull distributions are also used to represent various physical quantities, such as wind speed.

The Weibull distribution is a family of distributions that can assume the properties of several other distributions. For example, depending on the shape parameter you define, the Weibull distribution can be used to model the exponential and Rayleigh distributions, among others. The Weibull distribution is very flexible. When the Weibull shape parameter is equal to 1.0, the Weibull distribution is identical to the exponential distribution. The Weibull location parameter lets you set up an exponential distribution to start at a location other than 0.0. When the shape parameter is less than 1.0, the Weibull distribution becomes a steeply declining curve. A manufacturer might find this effect useful in describing part failures during a burn-in period.

The mathematical constructs for the Weibull distribution are as follows:

$$f(x) = \frac{\alpha}{\beta}\left[\frac{x}{\beta}\right]^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)^{\alpha}}$$

$$Mean = \beta\Gamma(1+\alpha^{-1})$$

$$Standard\ Deviation = \beta^2\left[\Gamma(1+2\alpha^{-1}) - \Gamma^2(1+\alpha^{-1})\right]$$

$$Skewness = \frac{2\Gamma^3(1+\beta^{-1}) - 3\Gamma(1+\beta^{-1})\Gamma(1+2\beta^{-1}) + \Gamma(1+3\beta^{-1})}{\left[\Gamma(1+2\beta^{-1}) - \Gamma^2(1+\beta^{-1})\right]^{3/2}}$$

$$Excess\ Kurtosis =$$

$$\frac{-6\Gamma^4(1+\beta^{-1}) + 12\Gamma^2(1+\beta^{-1})\Gamma(1+2\beta^{-1}) - 3\Gamma^2(1+2\beta^{-1}) - 4\Gamma(1+\beta^{-1})\Gamma(1+3\beta^{-1}) + \Gamma(1+4\beta^{-1})}{\left[\Gamma(1+2\beta^{-1}) - \Gamma^2(1+\beta^{-1})\right]^2}$$

Location (L), shape (α), and scale (β) are the distributional parameters, and Γ is the gamma function.

Input requirements:

Scale > 0 and can be any positive value
Shape ≥ 0.05
Location can take on any value

QUESTIONS

1. Why do you need to have profiles in a simulation?
2. Explain the differences between Pearson's product moment correlation coefficient and Spearman's rank-based correlation.
3. Will more or fewer trials be required to obtain: higher error levels, higher precision levels, and a wider confidence interval?
4. Explain the differences between error and precision and how these two concepts are linked.
5. If you know that two simulated variables are correlated but do not have the relevant correlation value, should you still go ahead and correlate them in a simulation?

Following are some hands-on exercises using Risk Simulator. The example files are located on *Start, Programs, Real Options Valuation, Risk Simulator, Examples*.

Basic Simulation Model

This sample model is used to illustrate how to use Risk Simulator for:

1. Running a Monte Carlo simulation
2. Viewing and interpreting forecast results
3. Setting seed values
4. Setting run preferences
5. Extracting simulation data
6. Creating new and switching among simulation profiles

Model Background

File Name: Basic Simulation Model.xls

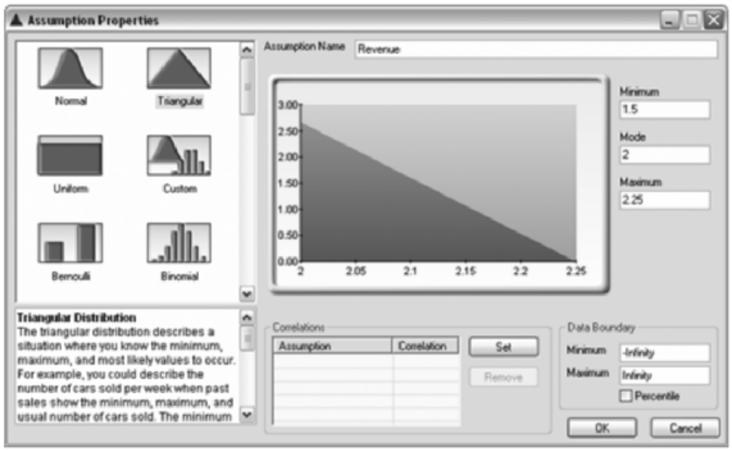
The model in the Static and Dynamic Model worksheet illustrates a very simple model with two input assumptions (revenue and cost) and an output forecast (income). The model on the left is a static model with single-point estimates while the model on the right is a dynamic model that we can create Monte Carlo assumptions and forecasts on. After running the simulation, the results can be extracted and further analyzed. Also, in this model, we can learn to set different simulation preferences, run a simulation with error and precision controls, as well as setting seed values. The model looks like the following, with a static unsimulated version on the left and a dynamic or stochastic version on the right:

STATIC MODEL			DYNAMIC MODEL			
Revenue	$	2.00	Revenue	$	2.00	<<---This is an Input Assumption
Cost	$	1.00	Cost	$	1.00	<<---This is an Input Assumption
Income	$	1.00	Income	$	1.00	<<---This is an Output Forecast

Running a Monte Carlo Simulation

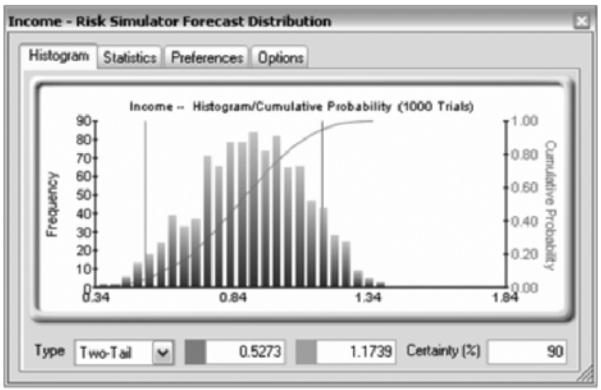
To run this model, simply:

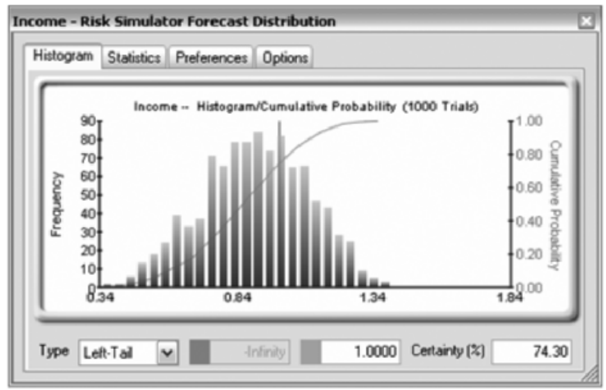
1. Select **Simulation | New Simulation Profile** (or click on the **New Profile icon**) and provide it a name.
2. Select cell *G8* and click on **Simulation | Set Input Assumption** (or click on the **Set Input Assumption icon**).
3. Select *Triangular Distribution* and set the *Min* = *1.50*, *Most Likely* = *2.00*, and *Max* = *2.25* and hit **OK**.
4. Select cell *G9* and set another input assumption. This time use *Uniform Distribution* with *Min* = *0.85* and *Max* = *1.25*.
5. Select cell *G10* and set that cell as the output forecast by clicking on **Simulation | Set Output Forecast**.
6. Select **Simulation | Run Simulation** (or click on the **Run icon**) to start the simulation.



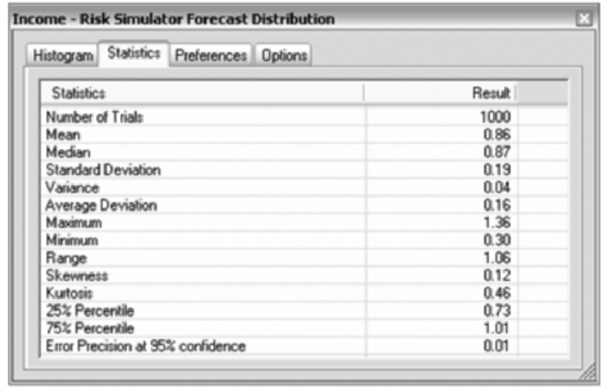
Viewing and Interpreting Forecast Results

The forecast chart is shown when the simulation is running. Once simulation is completed, the forecast chart can be used. The forecast chart has several tabs including the Histogram, Statistics, Preferences, and Options tabs. Of particular interest are the first two tabs. For instance, the first tab shows the output forecast's probability distribution in the form of a histogram, where the specific values can be determined using the certainty boxes. For instance, select Two-Tail and enter 90 in the certainty box and hit tab. The 90% confidence interval is shown (0.5273 and 1.1739). This means that there is a 5% chance that the income will fall below \$0.5273 and another 5% chance that it will be above \$1.1739. Alternatively, you can select Left Tail and enter 1.0 on the input box, hit tab, and see that the left-tail certainty is 74.30%, indicating that there is a 74.30% chance that the income will fall below \$1.0 (alternatively, there is a 25.70% chance that income will exceed \$1.0).





The statistics tab illustrates the statistical results of the forecast variable.



Note that your results will not be exactly those illustrated here. This is because a simulation (random number generation) was run, and by definition, the results will not be exactly the same every time. However, if a seed value is set (see next section), the results will be identical in every single run. This is important especially when you wish to obtain the same values in each simulation run (e.g., you need the live model to return the same results as a printed report during a presentation).

Setting a Seed Value

1. Reset the simulation by selecting **Simulation | Reset Simulation**.
2. Select **Simulation | Edit Simulation Profile**.
3. Select the *check box* for random number sequence.
4. Enter in a seed value (e.g., *999*) and hit **OK**.
5. **Run** the simulation and verify that the results are the same as the results obtained earlier.

Note that the random number sequence or seed number has to be a positive integer value. Running the same model with the same assumptions and forecasts with the identical seed value and same number of trials will always yield the same results. The number of simulation trials to run can be set in the same run properties box.

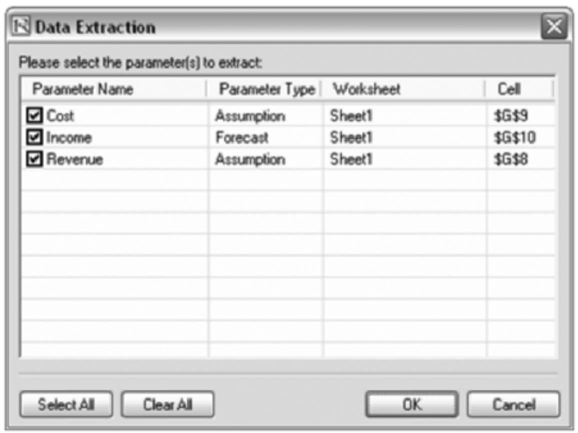
Setting Run Preferences

The run preferences dialog box allows you to specify the number of trials to run in a particular simulation. In addition, the simulation can also be pause if a computational error in Excel is encountered (e.g., #NUM or #ERROR). Correlations can also be specified between pairs of input assumptions and if Turn on Correlations is selected, these specified correlations will be imputed in the simulation.

Extracting Simulation Data

The simulation's assumptions and forecast data are stored in memory until the simulation is reset or when Excel is closed. If required, these raw data can be extracted into a separate Excel spreadsheet. To extract the data, simply:

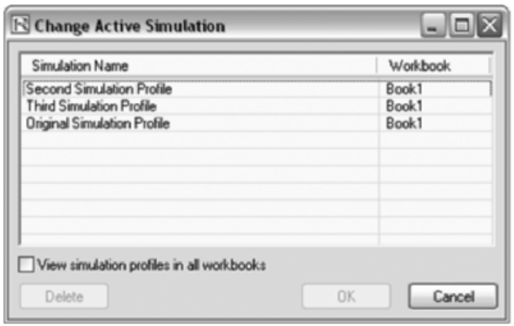
1. **Run** the simulation
2. After the simulation is completed select **Simulation | Tools | Extract Data**.
3. *Choose the relevant assumptions or forecasts* to extract and click **OK**.



Creating New and Switching Among Simulation Profiles

The same model can have multiple simulation profiles in Risk Simulator. That is, different users of the same model can in fact create his or her own simulation assumptions, forecasts, run preferences, and so forth. All these preferences are stored in separate simulation profiles and each profile can be run independently. This is a powerful feature that allows multiple users to run the same model their own way, or for the same user to run the model under different simulation conditions, thereby allowing for scenario analysis on Monte Carlo simulation. To create different profiles and switch among different profiles, simply:

1. Create several new profiles by clicking on **Simulation | New Simulation Profile**.
2. Add the relevant assumptions, forecasts, or change the run preferences as desired in each simulation profile.
3. Switch among different profiles by clicking on **Simulation | Change Active Simulation**.



Correlated Simulation

This sample model is used to illustrate how to use Risk Simulator for:

1. Running a simulation with correlated variables
2. Extracting data and determining the final correlations

Model Background
File Name: Correlated Simulation.xls
This model illustrates how to set up basic correlated variables in a Monte Carlo simulation, extract the data, and confirm that the resulting correlations closely match the input correlations in the model. The sample model looks like the following, with 6 different simulation assumptions:

A	100.00	D	100.00
B	100.00	E	100.00
C	100.00	F	100.00

All variables are Normal (100, 10) and correlated at 0.5 to each other, and 5,000 trials are run with a seed value of 123456.

Correlated Model
Simply open the existing profile, run the simulation, extract the data, and compute a correlation matrix. The results for 5,000 trials are provided in this example. Note that the input pairwise correlations were all 0.50 and the resulting correlations computed from the data are very close to the required input correlations.

Sample Results

The following lists a sample extracted dataset of the 5,000-trial simulation and the resulting correlation matrix computed on the data.

A	B	C	D	E	F
93.75	86.07	77.84	86.76	96.26	89.85
91.36	93.30	95.84	102.04	86.40	91.88
102.29	90.24	102.00	108.01	104.83	103.81
104.64	90.19	97.91	89.48	99.95	105.91
86.79	89.54	80.49	70.12	89.42	87.11
97.32	100.99	110.57	92.41	107.92	91.80
105.90	102.93	101.27	105.66	95.94	106.76
84.60	77.91	84.40	75.55	88.60	85.29
79.74	85.90	87.52	86.16	81.27	93.13
106.69	93.61	98.69	91.62	93.50	97.38
97.70	91.11	100.40	116.27	99.43	100.63
98.40	98.83	93.95	99.96	90.52	90.85
92.06	78.97	101.03	98.26	82.84	94.65
100.64	81.43	96.64	96.10	112.09	95.39
87.45	64.69	82.16	74.13	78.89	85.41
104.38	103.73	96.50	112.12	98.01	120.22

Correlation Matrix

	A	B	C	D	E	F
A	1					
B	0.49	1				
C	0.52	0.49	1			
D	0.51	0.51	0.50	1		
E	0.52	0.51	0.51	0.51	1	
F	0.50	0.48	0.50	0.49	0.50	1
Average Correlation:			0.50			

Correlation Effects Model

This sample model is used to illustrate how to use Risk Simulator for:

1. Creating correlated simulations and comparing the results between correlated and uncorrelated models
2. Extracting and manually computing (verifying) the assumptions' correlations

Model Background

File Name: Correlation Effects Model.xls

This model illustrates the effects of correlated simulation versus uncorrelated simulation. That is, when a pair of simulated assumptions are not correlated, positively correlated, and negatively correlated. The results can sometimes be very different. In addition, the simulated assumptions' raw data are extracted after the simulation and manual computations of their pairwise correlations are performed and the results indicate that the correlations hold after the simulation.

Correlation Model

	Without Correlation	Positive Correlation	Negative Correlation
Price	$2.00	$2.00	$2.00
Quantity	1.00	1.00	1.00
Revenue	$2.00	$2.00	$2.00

Note: Prices are set as Triangular Distributions (1.8, 2.0, 2.2) while Quantity are set as Uniform Distributions (0.9, 1.1) with correlations set at 0.0, +0.8, -0.8 at 1,000 trials with seed value 123456.

Running a Monte Carlo Simulation

To run this model, simply:

1. Click on **Simulation** I **Change Simulation Profile**, then choose **Correlation Profile** and click OK.
2. Run the simulation by clicking on the **Run** icon or **Simulation** I **Run Simulation**.

Viewing and Interpreting Forecast Results

The resulting simulation statistics indicate that the negatively correlated variables provide a tighter or smaller standard deviation or risk level. The positively correlated variables provide the highest standard deviation, while the uncorrelated variables have a standard deviation that is in between these two sets of positive and negatively correlated simulations. This is because negative correlations provide a diversification effect on the variables and hence tend to make the standard deviation slightly smaller. Thus, one needs to make sure to input correlations when there indeed are correlations between variables otherwise this interacting effect will not be accounted for in the simulation.

Revenue No Correlation - Risk Simulator Forecast Distribution

Histogram | Statistics | Preferences | Options

Statistics	Result
Number of Trials	1000
Mean	2.0006
Median	1.9933
Standard Deviation	0.1409
Variance	0.0199
Average Deviation	0.1173
Maximum	2.3676
Minimum	1.6579
Range	0.7096
Skewness	0.1593
Kurtosis	-0.6460
25% Percentile	1.8930
75% Percentile	2.1064
Error Precision at 95% confidence	0.4365%

Revenue Negative Correlation - Risk Simulator Forecast Distribution

Histogram | Statistics | Preferences | Options

Statistics	Result
Number of Trials	1000
Mean	2.0003
Median	1.9995
Standard Deviation	0.0750
Variance	0.0056
Average Deviation	0.0611
Maximum	2.2569
Minimum	1.8010
Range	0.4559
Skewness	0.1377
Kurtosis	-0.3783
25% Percentile	1.9449
75% Percentile	2.0541
Error Precision at 95% confidence	0.2324%

Revenue Positive Correlation - Risk Simulator Forecast Distribution

Histogram | Statistics | Preferences | Options

Statistics	Result
Number of Trials	1000
Mean	1.9981
Median	1.9911
Standard Deviation	0.1891
Variance	0.0357
Average Deviation	0.1590
Maximum	2.4147
Minimum	1.6258
Range	0.7889
Skewness	0.0982
Kurtosis	-0.9079
25% Percentile	1.8485
75% Percentile	2.1474
Error Precision at 95% confidence	0.5864%

Discounted Cash Flow, Return on Investment and Volatility Estimates

This sample model is used to illustrate how to use Risk Simulator for:

1. Running a Monte Carlo simulation in a financial ROI analysis.
2. Incorporating various correlations among variables (time-series and cross-sectional relationships).
3. Calculating forward-looking volatility based on the PV Asset approach.

Model Background

File Name: DCF, ROI and Volatility.xls

This is a generic discounted cash flow (DCF) model showing the net present value (NPV), internal rate of return (IRR), and return on investment (ROI) calculations. This model can be adapted to any industry and is used to illustrate how a Monte Carlo simulation can be applied to actual investment decisions and capital budgeting. In addition, the project's or asset's volatility is also computed in this example model. Volatility is a measure of risk and is a key input into real options analysis.

Discounted Cash Flow Model

Base Year	*2005*	*Sum PV Net Benefits*	***$4,481.91***	*Discount Type*	Discrete End-of-Year Discounting
Start Year	*2005*	*Sum PV Investments*	***$1,634.22***		
Market Risk-Adjusted Discount Rate	*15.00%*	*Net Present Value*	***$2,847.69***	*Model*	Include Terminal Valuation
Private-Risk Discount Rate	*5.00%*	*Internal Rate of Return*	***54.64%***		
Terminal Period Growth Rate	*2.00%*	*Return on Investment*	***174.25%***		
Effective Tax Rate	*40.00%*	*Profitability Index*	***2.74***		

	2005	*2006*	*2007*	*2008*	*2009*	*2010*	*2011*	*2012*	*2013*	*2014*
Product A Avg Price/Unit	$10.00	$10.50	$11.00	$11.50	$12.00	$12.00	$12.00	$12.00	$12.00	$12.00
Product B Avg Price/Unit	$12.25	$12.50	$12.75	$13.00	$13.25	$13.25	$13.25	$13.25	$13.25	$13.25
Product C Avg Price/Unit	$15.15	$15.30	$15.45	$15.60	$15.75	$15.75	$15.75	$15.75	$15.75	$15.75
Product A Sale Quantity ('000s)	50	50	50	50	50	50	50	50	50	50
Product B Sale Quantity ('000s)	35	35	35	35	35	35	35	35	35	35
Product C Sale Quantity ('000s)	20	20	20	20	20	20	20	20	20	20
Total Revenues	***$1,231.75***	***$1,268.50***	***$1,305.25***	***$1,342.00***	***$1,378.75***	***$1,378.75***	***$1,378.75***	***$1,378.75***	***$1,378.75***	***$1,378.75***
Direct Cost of Goods Sold	$184.76	$190.28	$195.79	$201.30	$206.81	$206.81	$206.81	$206.81	$206.81	$206.81
Gross Profit	***$1,046.99***	***$1,078.23***	***$1,109.46***	***$1,140.70***	***$1,171.94***	***$1,171.94***	***$1,171.94***	***$1,171.94***	***$1,171.94***	***$1,171.94***
Operating Expenses	$157.50	$157.50	$157.50	$157.50	$157.50	$157.50	$157.50	$157.50	$157.50	$157.50
Sales, General and Admin. Costs	$15.75	$15.75	$15.75	$15.75	$15.75	$15.75	$15.75	$15.75	$15.75	$15.75
Operating Income (EBITDA)	***$873.74***	***$904.98***	***$936.21***	***$967.45***	***$998.69***	***$998.69***	***$998.69***	***$998.69***	***$998.69***	***$998.69***
Depreciation	$10.00	$10.00	$10.00	$10.00	$10.00	$10.00	$10.00	$10.00	$10.00	$10.00
Amortization	$3.00	$3.00	$3.00	$3.00	$3.00	$3.00	$3.00	$3.00	$3.00	$3.00
EBIT	***$860.74***	***$891.98***	***$923.21***	***$954.45***	***$985.69***	***$985.69***	***$985.69***	***$985.69***	***$985.69***	***$985.69***
Interest Payments	$2.00	$2.00	$2.00	$2.00	$2.00	$3.00	$4.00	$5.00	$6.00	$7.00
EBT	***$858.74***	***$889.98***	***$921.21***	***$952.45***	***$983.69***	***$982.69***	***$981.69***	***$980.69***	***$979.69***	***$978.69***
Taxes	$343.50	$355.99	$368.49	$380.98	$393.48	$393.08	$392.68	$392.28	$391.88	$391.48
Net Income	***$515.24***	***$533.99***	***$552.73***	***$571.47***	***$590.21***	***$589.61***	***$589.01***	***$588.41***	***$587.81***	***$587.21***

Note: The model is protected to prevent accidentally overriding key formulas. To unlock, click on Tools, Protection, Unprotect Sheet and enter *risksimulator* as the password. After unlocking the worksheet, the formulas will be visible (be careful not to override key formulas).

Running the Model

Create a new simulation profile in this model, and set several assumptions for the product price and quantity sold for each product line. You can either link the price and quantity time-series through an annual growth rate and hence only one assumption at the beginning of 2005 is required for each product line. In contrast, you can set assumptions on each product and price cell and set correlations between the years, creating a series of autocorrelated cash flows. Then, simply unlock the worksheet and set the NPV as a forecast and run the simulation.

CHAPTER 6

Pandora's Toolbox

This chapter deals with the Risk Simulator software's analytical tools. These analytical tools are discussed through example applications of the Risk Simulator software, complete with step-by-step illustrations. These tools are very valuable to analysts working in the realm of risk analysis. The applicability of each tool is discussed in detail in this chapter. All of the example files used in this chapter are found by going to *Start, Programs, Real Options Valuation, Risk Simulator, Examples*.

TORNADO AND SENSITIVITY TOOLS IN SIMULATION

Theory

One of the powerful simulation tools is tornado analysis—it captures the static impacts of each variable on the outcome of the model; that is, the tool automatically perturbs each variable in the model a preset amount, captures the fluctuation on the model's forecast or final result, and lists the resulting perturbations ranked from the most significant to the least. Figures 6.1 through 6.6 illustrate the application of a tornado analysis. For instance, Figure 6.1 is a sample discounted cash-flow model where the input assumptions in the model are shown. The question is, what are the critical success drivers that affect the model's output the most? That is, what really drives the net present value of $96.63 or which input variable impacts this value the most?

The tornado chart tool can be obtained through *Simulation | Tools | Tornado Analysis*. To follow along the first example, open the *Tornado and Sensitivity Charts (Linear)* file in the examples folder. Figure 6.2 shows this sample model where cell G6 containing the net present value is chosen as the target result to be analyzed. The target cell's precedents in the model are used in creating the tornado chart. Precedents are all the input and intermediate variables that affect the outcome of the model. For instance, if the model consists of $A = B + C$, and where $C = D + E$, then B, D, and E are the precedents for A (C is not a precedent as it is only an intermediate calculated

Discounted Cash Flow Model

Base Year	*2005*	*Sum PV Net Benefits*	*$1,896.63*
Market Risk-Adjusted Discount Rate	*15.00%*	*Sum PV Investments*	*$1,800.00*
Private-Risk Discount Rate	*5.00%*	*Net Present Value*	*$96.63*
Annualized Sales Growth Rate	*2.00%*	*Internal Rate of Return*	*18.80%*
Price Erosion Rate	*5.00%*	*Return on Investment*	*5.37%*
Effective Tax Rate	*40.00%*		

	2005	*2006*	*2007*	*2008*	*2009*
Product A Avg Price/Unit	$10.00	$9.50	$9.03	$8.57	$8.15
Product B Avg Price/Unit	$12.25	$11.64	$11.06	$10.50	$9.98
Product C Avg Price/Unit	$15.15	$14.39	$13.67	$12.99	$12.34
Product A Sale Quantity ('000s)	50.00	51.00	52.02	53.06	54.12
Product B Sale Quantity ('000s)	35.00	35.70	36.41	37.14	37.89
Product C Sale Quantity ('000s)	20.00	20.40	20.81	21.22	21.65
Total Revenues	*$1,231.75*	*$1,193.57*	*$1,156.57*	*$1,120.71*	*$1,085.97*
Direct Cost of Goods Sold	$184.76	$179.03	$173.48	$168.11	$162.90
Gross Profit	*$1,046.99*	*$1,014.53*	*$983.08*	*$952.60*	*$923.07*
Operating Expenses	$157.50	$160.65	$163.86	$167.14	$170.48
Sales, General and Admin. Costs	$15.75	$16.07	$16.39	$16.71	$17.05
Operating Income (EBITDA)	*$873.74*	*$837.82*	*$802.83*	*$768.75*	*$735.54*
Depreciation	$10.00	$10.00	$10.00	$10.00	$10.00
Amortization	$3.00	$3.00	$3.00	$3.00	$3.00
EBIT	*$860.74*	*$824.82*	*$789.83*	*$755.75*	*$722.54*
Interest Payments	$2.00	$2.00	$2.00	$2.00	$2.00
EBT	*$858.74*	*$822.82*	*$787.83*	*$753.75*	*$720.54*
Taxes	$343.50	$329.13	$315.13	$301.50	$288.22
Net Income	*$515.24*	*$493.69*	*$472.70*	*$452.25*	*$432.33*
Noncash Depreciation Amortization	$13.00	$13.00	$13.00	$13.00	$13.00
Noncash: Change in Net Working Capital	$0.00	$0.00	$0.00	$0.00	$0.00
Noncash: Capital Expenditures	$0.00	$0.00	$0.00	$0.00	$0.00
Free Cash Flow	*$528.24*	*$506.69*	*$485.70*	*$465.25*	*$445.33*
Investment Outlay	*$1,800.00*				

Financial Analysis					
Present Value of Free Cash Flow	*$528.24*	*$440.60*	*$367.26*	*$305.91*	*$254.62*
Present Value of Investment Outlay	*$1,800.00*	*$0.00*	*$0.00*	*$0.00*	*$0.00*
Net Cash Flows	($1,271.76)	$506.69	$485.70	$465.25	$445.33

FIGURE 6.1 Sample discounted cash flow model.

value). Figure 6.2 also shows the testing range of each precedent variable used to estimate the target result. If the precedent variables are simple inputs, then the testing range will be a simple perturbation based on the range chosen (e.g., the default is ±10 percent). Each precedent variable can be perturbed at different percentages if required. A wider range is important as it is better able to test extreme values rather than smaller perturbations around the expected values. In certain circumstances, extreme values may have a larger, smaller, or unbalanced impact (e.g., nonlinearities may occur where increasing or decreasing economies of scale and scope creep in for larger or

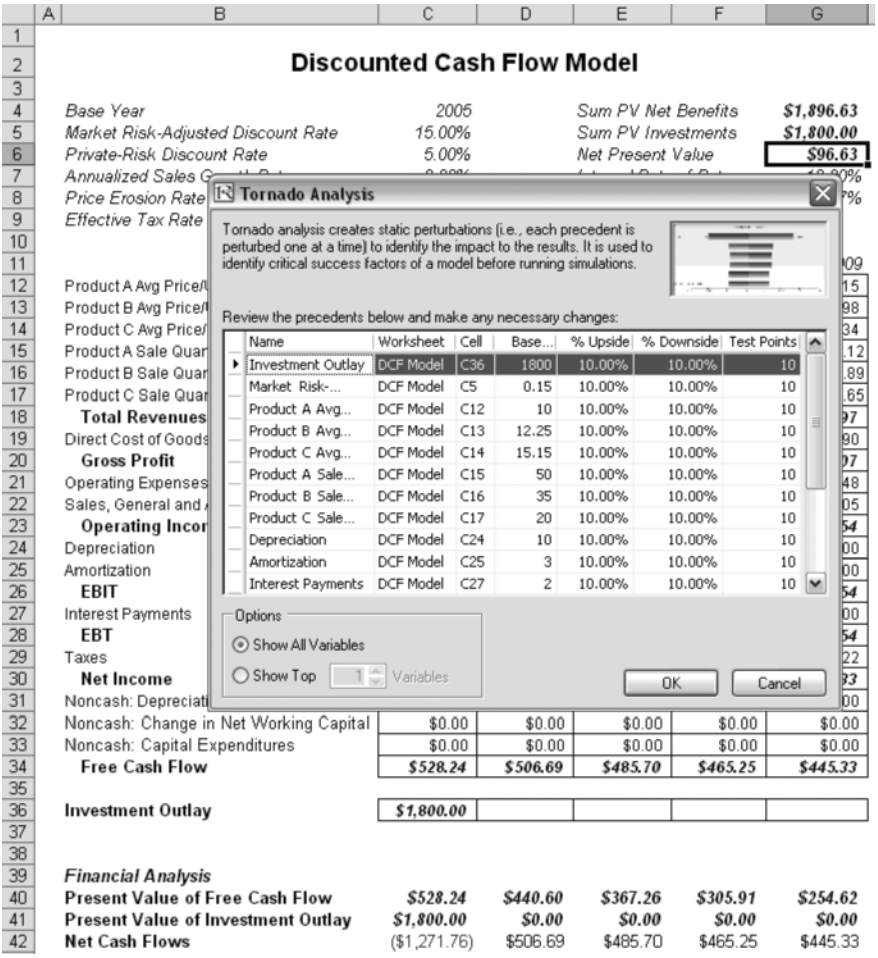


FIGURE 6.2 Running tornado analysis.

smaller values of a variable) and only a wider range will capture this nonlinear impact.

Procedure

Use the following steps to create a tornado analysis:

1. Select the single output cell (i.e., a cell with a function or equation) in an Excel model (e.g., cell G6 is selected in our example).
2. Select *Simulation* | *Tools* | *Tornado Analysis*.

3. Review the precedents and rename them as appropriate (renaming the precedents to shorter names allows a more visually pleasing tornado and spider chart) and click OK. Alternatively, click on *Use Cell Address* to apply cell locations as the variable names.

Results Interpretation

Figure 6.3 shows the resulting tornado analysis report, which indicates that capital investment has the largest impact on net present value (NPV), followed by tax rate, average sale price and quantity demanded of the product lines, and so forth. The report contains four distinct elements:

1. Statistical summary listing the procedure performed.
2. A sensitivity table (Table 6.1) shows the starting NPV base value of \$96.63 and how each input is changed (e.g., investment is changed from \$1,800 to \$1,980 on the upside with a +10 percent swing, and from \$1,800 to \$1,620 on the downside with a –10 percent swing). The resulting upside and downside values on NPV are –\$83.37 and \$276.63, with a total change of \$360, making it the variable with the highest impact on NPV. The precedent variables are ranked from the highest impact to the lowest impact.
3. The spider chart (Figure 6.4) illustrates these effects graphically. The y-axis is the NPV target value whereas the x-axis depicts the percentage change on each of the precedent values (the central point is the base case value at \$96.63 at 0 percent change from the base value of each precedent). Positively sloped lines indicate a positive relationship or effect while negatively sloped lines indicate a negative relationship (e.g., investment is negatively sloped, which means that the higher the investment level, the lower the NPV). The absolute value of the slope indicates the magnitude of the effect computed as the percentage change in the result given a percentage change in the precedent (a steep line indicates a higher impact on the NPV y-axis given a change in the precedent x-axis).
4. The tornado chart (Figure 6.5) illustrates the results in another graphical manner, where the highest impacting precedent is listed first. The x-axis is the NPV value with the center of the chart being the base case condition. Green (lighter) bars in the chart indicate a positive effect while red (darker) bars indicate a negative effect. Therefore, for investments, the red (darker) bar on the right side indicates a negative effect of investment on higher NPV—in other words, capital investment and NPV are negatively correlated. The opposite is true for price and quantity of products A to C (their green or lighter bars are on the right side of the chart).

Statistical Summary

One of the powerful simulation tools is the tornado chart—it captures the static impacts of each variable on the outcome of the model. That is, the tool automatically perturbs each precedent variable in the model a user-specified preset amount, captures the fluctuation on the model's forecast or final result, and lists the resulting perturbations ranked from the most significant to the least. Precedents are all the input and intermediate variables that affect the outcome of the model. For instance, if the model consists of $A = B + C$, where $C = D + E$, then B, D, and E are the precedents for A (C is not a precedent as it is only an intermediate calculated value). The range and number of values perturbed is user-specified and can be set to test extreme values rather than smaller perturbations around the expected values. In certain circumstances, extreme values may have a larger, smaller, or unbalanced impact (e.g., nonlinearities may occur where increasing or decreasing economies of scale and scope creep occurs for larger or smaller values of a variable) and only a wider range will capture this nonlinear impact.

A tornado chart lists all the inputs that drive the model, starting from the input variable that has the most effect on the results. The chart is obtained by perturbing each precedent input at some consistent range (e.g., ± 10% from the base case) one at a time, and comparing their results to the base case. A spider chart looks like a spider with a central body and its many legs protruding. The positively sloped lines indicate a positive relationship, while a negatively sloped line indicates a negative relationship. Further, spider charts can be used to visualize linear and nonlinear relationships. The tornado and spider charts help identify the critical success factors of an output cell in order to identify the inputs to simulate. The identified critical variables that are uncertain are the ones that should be simulated. Do not waste time simulating variables that are neither uncertain nor have little impact on the results.

Result

	Base Value: 96.6261638553219			Input Changes		
Precedent Cell	Output Downside	Output Upside	Effective Range	Input Downside	Input Upside	Base Case Value
Investment	$276.63	($83.37)	360.00	$1,620.00	$1,980.00	$1,800.00
Tax Rate	$219.73	($26.47)	246.20	36.00%	44.00%	40.00%
A Price	$3.43	$189.83	186.40	$9.00	$11.00	$10.00
B Price	$16.71	$176.55	159.84	$11.03	$13.48	$12.25
A Quantity	$23.18	$170.07	146.90	45.00	55.00	50.00
B Quantity	$30.53	$162.72	132.19	31.50	38.50	35.00
C Price	$40.15	$153.11	112.96	$13.64	$16.67	$15.15
C Quantity	$48.05	$145.20	97.16	18.00	22.00	20.00
Discount Rate	$138.24	$57.03	81.21	13.50%	16.50%	15.00%
Price Erosion	$116.80	$76.64	40.16	4.50%	5.50%	5.00%
Sales Growth	$90.59	$102.69	12.10	1.80%	2.20%	2.00%
Depreciation	$95.08	$98.17	3.08	$9.00	$11.00	$10.00
Interest	$97.09	$96.16	0.93	$1.80	$2.20	$2.00
Amortization	$96.16	$97.09	0.93	$2.70	$3.30	$3.00
Capex	$96.63	$96.63	0.00	$0.00	$0.00	$0.00
Net Capital	$96.63	$96.63	0.00	$0.00	$0.00	$0.00

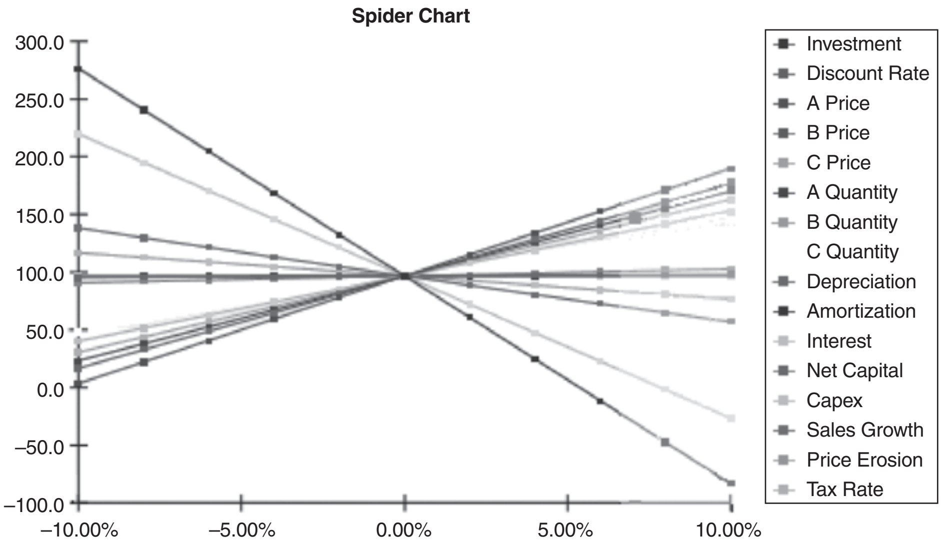


FIGURE 6.3 Tornado analysis report.

TABLE 6.1 Sensitivity Table

	Base Value: 96.6261638553219			Input Changes		
Precedent Cell	**Output Downside**	**Output Upside**	**Effective Range**	**Input Downside**	**Input Upside**	**Base Case Value**
Investment	$276.63	($83.37)	360.00	$1,620.00	$1,980.00	$1,800.00
Tax Rate	$219.73	($26.47)	246.20	36.00%	44.00%	40.00%
A Price	$3.43	$189.83	186.40	$9.00	$11.00	$10.00
B Price	$16.71	$176.55	159.84	$11.03	$13.48	$12.25
A Quantity	$23.18	$170.07	146.90	45.00	55.00	50.00
B Quantity	$30.53	$162.72	132.19	31.50	38.50	35.00
C Price	$40.15	$153.11	112.96	$13.64	$16.67	$15.15
C Quantity	$48.05	$145.20	97.16	18.00	22.00	20.00
Discount Rate	$138.24	$57.03	81.21	13.50%	16.50%	15.00%
Price Erosion	$116.80	$76.64	40.16	4.50%	5.50%	5.00%
Sales Growth	$90.59	$102.69	12.10	1.80%	2.20%	2.00%
Depreciation	$95.08	$98.17	3.08	$9.00	$11.00	$10.00
Interest	$97.09	$96.16	0.93	$1.80	$2.20	$2.00
Amortization	$96.16	$97.09	0.93	$2.70	$3.30	$3.00
Capex	$96.63	$96.63	0.00	$0.00	$0.00	$0.00
Net Capital	$96.63	$96.63	0.00	$0.00	$0.00	$0.00

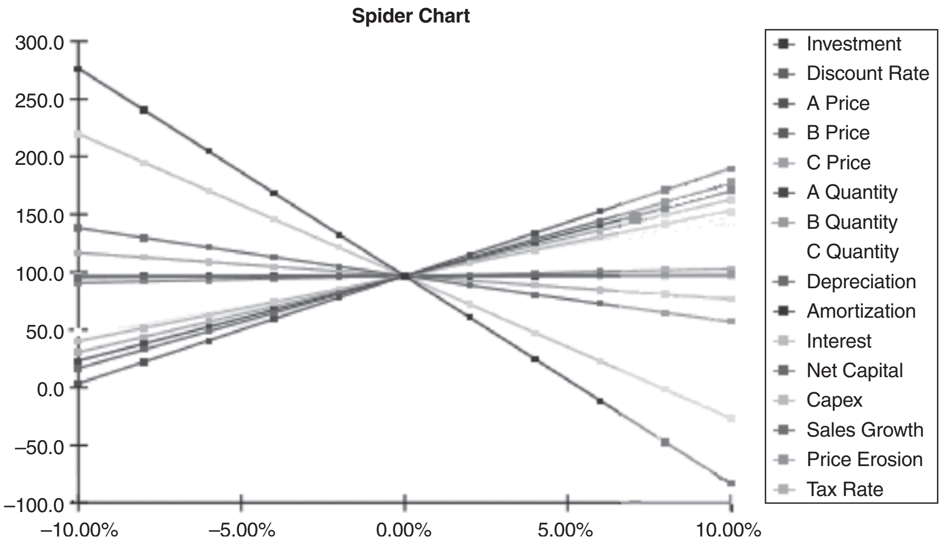


FIGURE 6.4 Spider chart.

Notes

Remember that tornado analysis is a *static* sensitivity analysis applied on each input variable in the model; that is, each variable is perturbed individually and the resulting effects are tabulated. This makes tornado analysis a key component to execute before running a simulation. One of the very first steps in risk analysis is where the most important impact drivers in the-model are captured and identified. The next step is to identify which of these important impact drivers are uncertain. These uncertain impact drivers are the critical success drivers of a project, where the results of the model depend on these critical success drivers. These variables are the ones that should be simulated. Do not waste time simulating variables that are neither uncertain nor have little impact on the results. Tornado charts assist in identifying these critical success drivers quickly and easily. Following this example, it might be that price and quantity should be simulated, assuming if the required investment and effective tax rate are both known in advance and unchanging.

Although the tornado chart is easier to read, the spider chart is important to determine if there are any nonlinearities in the model. For instance, Figure 6.6 shows another spider chart where nonlinearities are fairly evident (the lines on the graph are not straight but curved). The example model used is *Tornado and Sensitivity Charts (Nonlinear)*, which applies the Black–Scholes option pricing model. Such nonlinearities cannot be ascer-

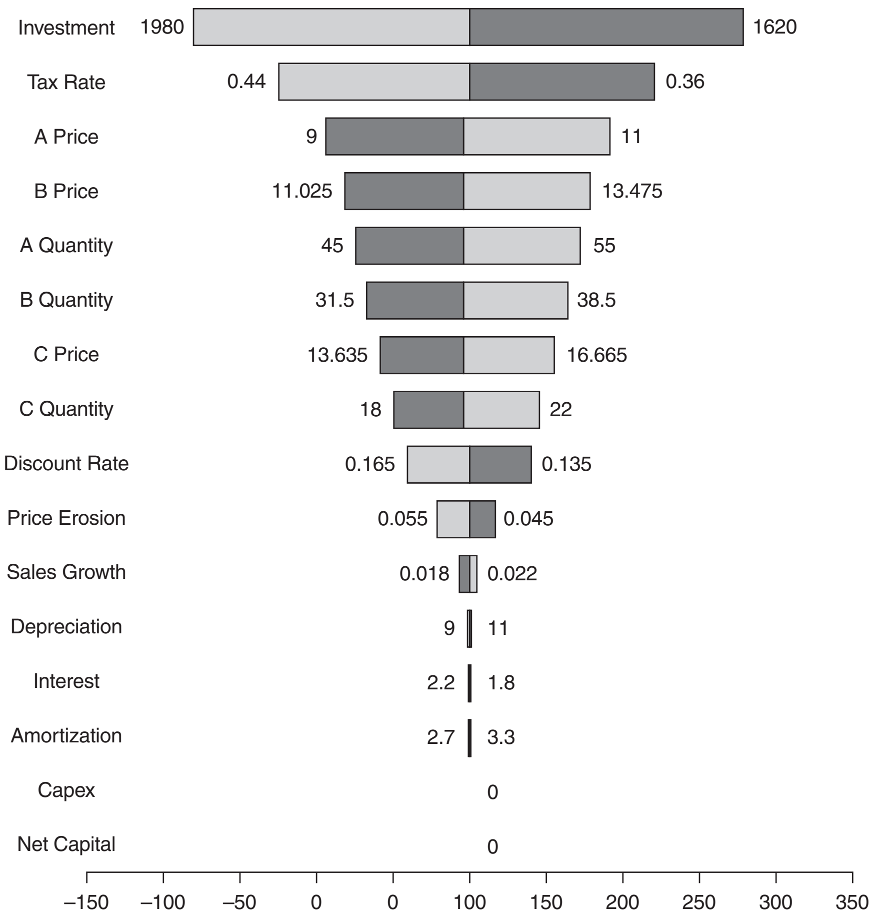


FIGURE 6.5 Tornado chart.

tained from a tornado chart and may be important information in the model or provide decision makers important insight into the model's dynamics. For instance, in this Black–Scholes model, the fact that stock price and strike price are nonlinearly related to the option value is important to know. This characteristic implies that option value will not increase or decrease proportionally to the changes in stock or strike price, and that there might be some interactions between these two prices as well as other variables. As another example, an engineering model depicting nonlinearities might indicate that a particular part or component, when subjected to a high enough force or tension, will break. Clearly, it is important to understand such nonlinearities.

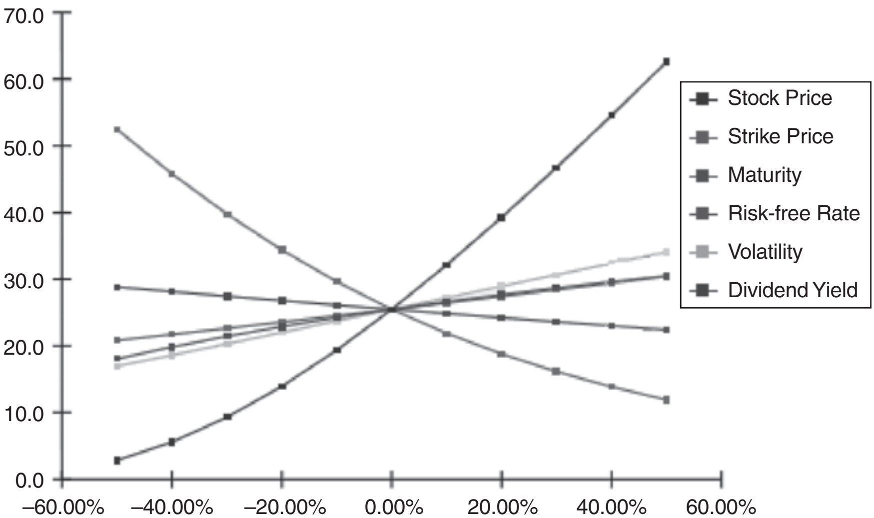


FIGURE 6.6 Nonlinear spider chart.

SENSITIVITY ANALYSIS

Theory

A related feature is sensitivity analysis. While tornado analysis (tornado charts and spider charts) applies static perturbations *before* a simulation run, sensitivity analysis applies dynamic perturbations created *after* the simulation run. Tornado and spider charts are the results of static perturbations, meaning that each precedent or assumption variable is perturbed a preset amount one at a time, and the fluctuations in the results are tabulated. In contrast, sensitivity charts are the results of dynamic perturbations in the sense that multiple assumptions are perturbed simultaneously and their interactions in the model and correlations among variables are captured in the fluctuations of the results. Tornado charts therefore identify which variables drive the results the most and hence are suitable for simulation, whereas sensitivity charts identify the impact to the results when multiple interacting variables are simulated together in the model. This effect is clearly illustrated in Figure 6.7. Notice that the ranking of critical success drivers is similar to the tornado chart in the previous examples. However, if correlations are added between the assumptions, Figure 6.8 shows a very different picture. Notice, for instance, price erosion had little impact on NPV, but when some of the input assumptions are correlated, the interaction that exists between these correlated variables makes price erosion have more impact. Note that

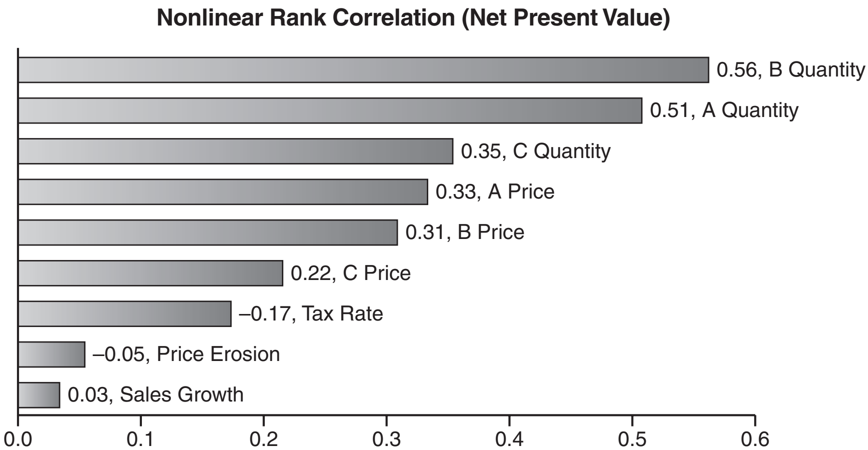


FIGURE 6.7 Sensitivity chart without correlations.

tornado analysis cannot capture these correlated dynamic relationships. Only after a simulation is run will such relationships become evident in a sensitivity analysis. A tornado chart's presimulation critical success factors will therefore sometimes be different from a sensitivity chart's postsimulation critical success factor. The postsimulation critical success factors should be the ones that are of interest as these more readily capture the model precedents' interactions.

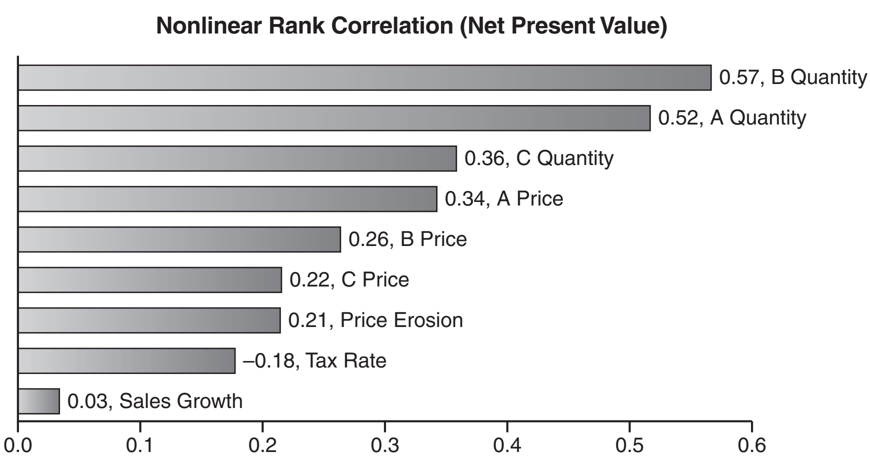


FIGURE 6.8 Sensitivity chart with correlations.

Procedure

Use the following steps to create a sensitivity analysis:

1. Open or create a model, define assumptions and forecasts, and run the simulation—the example here uses the Tornado and Sensitivity Charts (Linear) file.
2. Select *Simulation | Tools | Sensitivity Analysis*.
3. Select the forecast of choice to analyze and click OK (Figure 6.9).

Note that sensitivity analysis cannot be run unless assumptions and forecasts have been defined, and a simulation has been run.

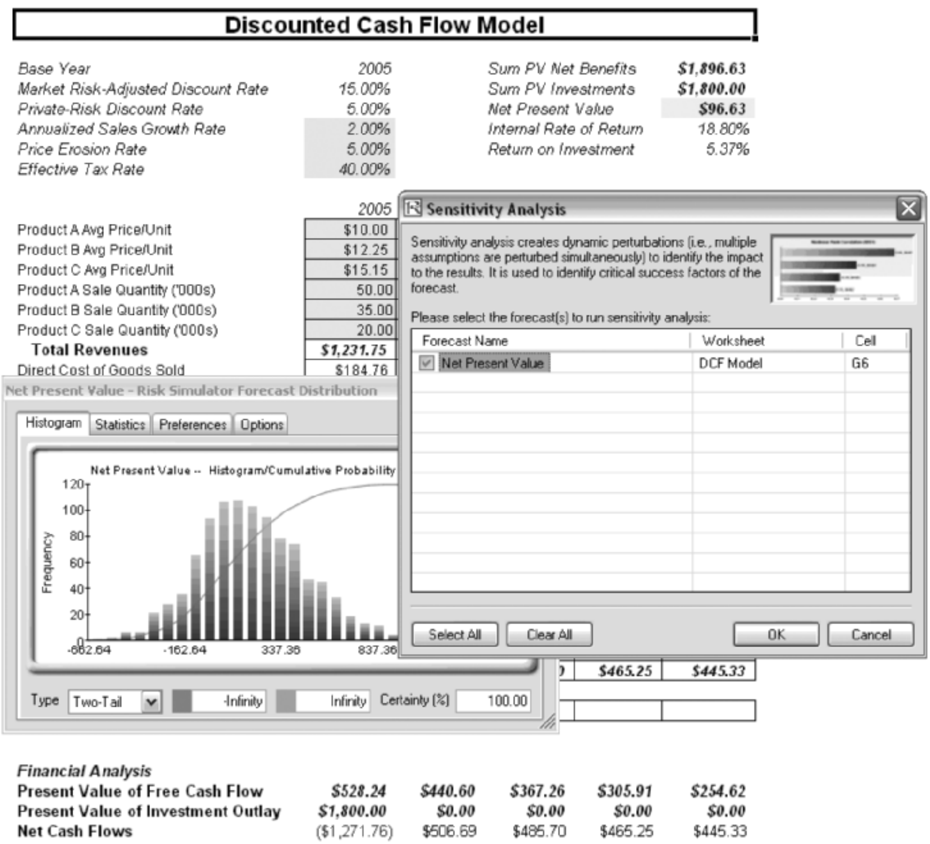


FIGURE 6.9 Running sensitivity analysis.

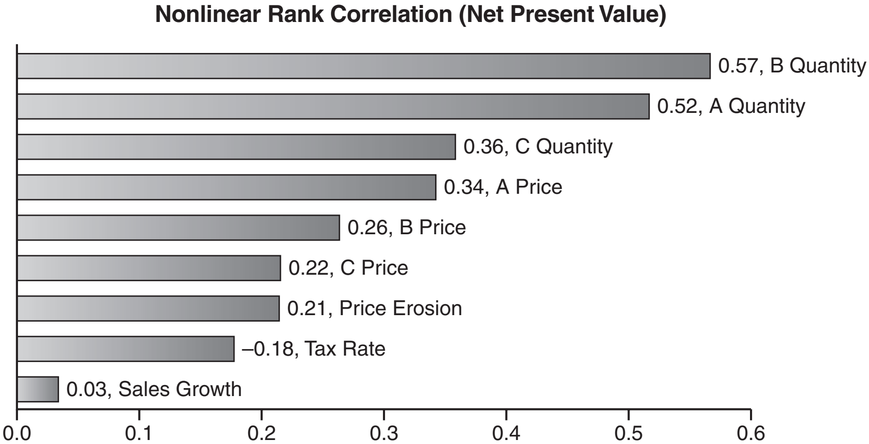


FIGURE 6.10 Rank correlation chart.

Results Interpretation

The results of the sensitivity analysis comprise a report and two key charts. The first is a nonlinear rank correlation chart (Figure 6.10) that ranks from highest to lowest the assumption–forecast correlation pairs. These correlations are nonlinear and nonparametric, making them free of any distributional requirements (i.e., an assumption with a Weibull distribution can be compared to another with a Beta distribution). The results from this chart are fairly similar to that of the tornado analysis seen previously (of course without the capital investment value, which we decided was a known value and hence was not simulated), with one special exception. Tax rate was relegated to a much lower position in the sensitivity analysis chart (Figure 6.10) as compared to the tornado chart (Figure 6.5). This is because by itself, tax rate will have a significant impact, but once the other variables are interacting in the model, it appears that tax rate has less of a dominant effect (because tax rate has a smaller distribution as historical tax rates tend not to fluctuate too much, and also because tax rate is a straight percentage value of the income before taxes, where other precedent variables have a larger effect on NPV). This example proves that performing sensitivity analysis after a simulation run is important to ascertain if there are any interactions in the model and if the effects of certain variables still hold. The second chart (Figure 6.11) illustrates the percent variation explained; that is, of the fluctuations in the forecast, how much of the variation can be explained by each of the assumptions after accounting for all the interactions among

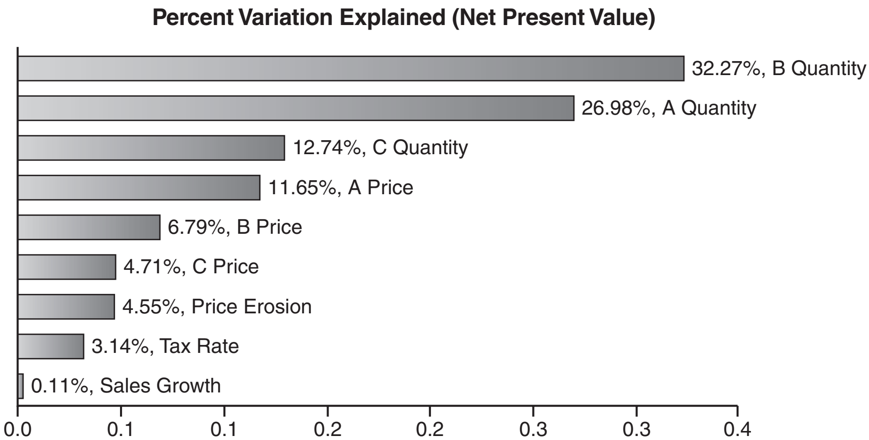


FIGURE 6.11 Contribution to variance chart.

variables? Notice that the sum of all variations explained is usually close to 100 percent (sometimes other elements impact the model, but they cannot be captured here directly), and if correlations exist, the sum may sometimes exceed 100 percent (due to the interaction effects that are cumulative).

Notes

Tornado analysis is performed before a simulation run while sensitivity analysis is performed after a simulation run. Spider charts in tornado analysis can consider nonlinearities while rank correlation charts in sensitivity analysis can account for nonlinear and distributional-free conditions.

DISTRIBUTIONAL FITTING: SINGLE VARIABLE AND MULTIPLE VARIABLES

Theory

Another powerful simulation tool is distributional fitting; that is, which distribution does an analyst or engineer use for a particular input variable in a model? What are the relevant distributional parameters? If no historical data exist, then the analyst must make assumptions about the variables in question. One approach is to use the Delphi method, where a group of experts are tasked with estimating the behavior of each variable. For instance, a group of mechanical engineers can be tasked with evaluating the extreme possibilities of a spring coil's diameter through rigorous experimentation or guesstimates. These values can be used as the variable's input parameters (e.g., uniform distribution with extreme values between 0.5 and 1.2). When testing is not possible (e.g., market share and revenue growth rate), management can still make estimates of potential outcomes and provide the best-case, most-likely case, and worst-case scenarios, whereupon a triangular or custom distribution can be created.

However, if reliable historical data are available, distributional fitting can be accomplished. Assuming that historical patterns hold and that history tends to repeat itself, then historical data can be used to find the best-fitting distribution with their relevant parameters to better define the variables to be simulated. Figure 6.12, Figure 6.13, and Figure 6.14 illustrate a distributional-fitting example. The following illustration uses the *Data Fitting* file in the examples folder.

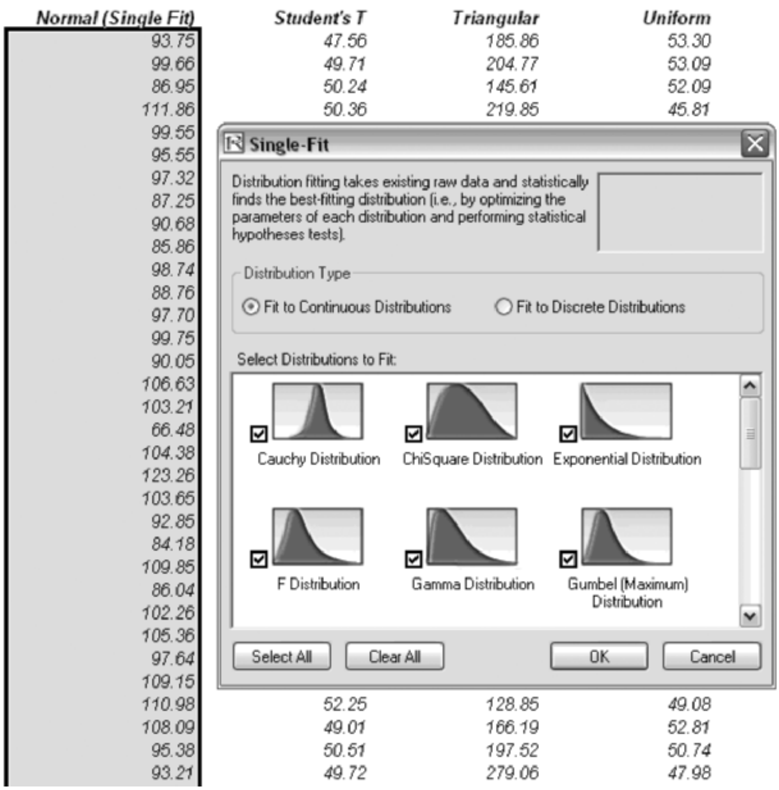


FIGURE 6.12 Single-variable distributional fitting.

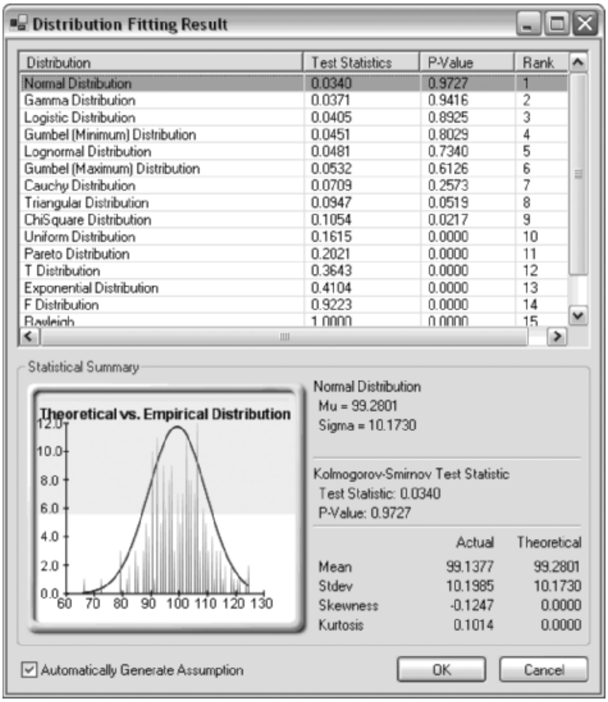


FIGURE 6.13 Distributional fitting result.

Procedure

Use the following steps to perform a distribution fitting model:

1. Open a spreadsheet with existing data for fitting (e.g., use the Data Fitting example file).
2. Select the data you wish to fit not including the variable name (data should be in a single column with multiple rows).
3. Select *Simulation | Tools | Distributional Fitting (Single-Variable)*.
4. Select the specific distributions you wish to fit to or keep the default where all distributions are selected and click OK (Figure 6.12).
5. Review the results of the fit, choose the relevant distribution you want, and click OK (Figure 6.13).

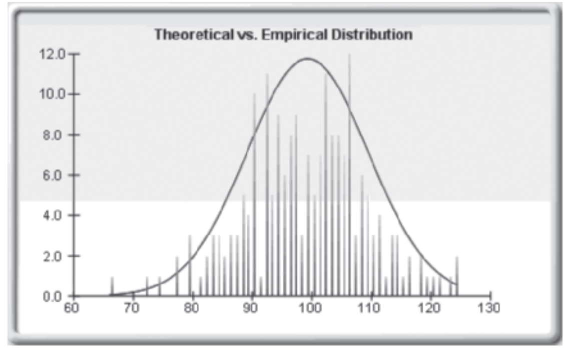
Results Interpretation

The null hypothesis (H_o) being tested is such that the fitted distribution is the same distribution as the population from which the sample data to be fitted comes. Thus, if the computed p-value is lower than a critical alpha level

Single Variable Distributional Fitting

Statistical Summary

Fitted Assumption	99.14	
Fitted Distribution	Normal Distribution	
Mu	99.28	
Sigma	10.17	
Kolmogorov-Smimov Statistic	0.03	
P-Value for Test Statistic	0.9727	
	Actual	Theoretical
Mean	99.28	99.28
Standard Deviation	10.17	10.17
Skewness	0.00	0.00
Excess Kurtosis	0.00	0.00



Original Fitted Data

93.75	99.66	86.95	11.86	99.55	95.55	97.32	87.25	90.68	85.86	98.74	88.76	97.70
99.75	90.05	106.63	103.21	66.48	104.38	123.26	103.65	92.85	84.18	109.85	86.04	102.26
105.36	97.64	109.15	110.98	108.09	95.38	93.21	83.86	100.17	110.17	103.72	120.52	95.09
115.18	83.64	90.23	92.44	92.37	92.70	110.81	72.67	104.23	96.47	121.15	94.92	77.26
103.45	96.75	93.91	101.91	134.14	90.95	107.13	92.02	96.43	96.35	88.30	108.48	113.50
101.40	104.72	102.43	113.59	124.15	109.24	105.34	104.57	97.83	94.39	116.19	84.66	101.17
106.13	107.17	95.83	106.67	92.42	79.64	94.15	106.00	113.45	92.63	94.51	93.05	96.19
100.85	83.34	111.82	118.12	87.17	103.66	106.93	82.45	102.74	86.82	106.68	89.61	94.56
101.34	91.32	102.02	82.51	104.46	84.72	105.05	108.40	106.59	109.43	92.49	94.52	94.00
105.92	88.13	96.41	101.45	79.93	89.68	102.91	114.95	92.58	94.05	107.90	111.05	90.58
97.09	105.44	94.95	102.55	77.41	108.53	90.54	100.41	106.83	99.63	79.71	89.32	116.30
98.27	101.73	90.84	74.45	102.24	103.34	96.51	114.55	93.94	106.29	102.95	112.73	98.09
108.20	105.80	106.48	102.88	104.93	103.00	99.10	108.52	101.31	88.17	90.62	96.53	106.03
109.12	104.23	90.34	95.12	102.03	100.00	118.17	99.06	81.89	104.29	92.68	114.69	102.49
119.21	106.20	88.26	92.45	105.15	103.79	100.84	95.19	85.10	97.25	87.65	97.58	111.44
99.52	89.83	97.86	90.96	97.14								

FIGURE 6.14 Distributional fitting report.

(typically 0.10 or 0.05), then the distribution is the wrong distribution. Conversely, the *higher the p-value, the better the distribution fits the data*. Roughly, you can think of p-value as a *percentage explained*; that is, if the p-value is 0.9727 (Figure 6.13), then setting a normal distribution with a mean of 99.28 and a standard deviation of 10.17 explains about 97.27 percent of the variation in the data, indicating an especially good fit. The data was from a 1,000-trial simulation in Risk Simulator based on a normal distribution with a mean of 100 and a standard deviation of 10. Because only 1,000 trials were simulated, the resulting distribution is fairly close to the specified distributional parameters, and in this case, about a 97.27 percent precision.

Both the results (Figure 6.13) and the report (Figure 6.14) show the test statistic, p-value, theoretical statistics (based on the selected distribution), empirical statistics (based on the raw data), the original data (to maintain a record of the data used), and the assumption complete with the relevant distributional parameters (i.e., if you selected the option to automatically generate assumption and if a simulation profile already exists). The results also rank all the selected distributions and how well they fit the data.

Fitting Multiple Variables

For fitting multiple variables, the process is fairly similar to fitting individual variables. However, the data should be arranged in columns (i.e., each variable is arranged as a column) and all the variables are fitted. The same analysis is performed when fitting multiple variables as when single variables are fitted. The difference here is that only the final report will be generated and you do not get to review each variable's distributional rankings. If the rankings are important, run the single-variable fitting procedure instead, on one variable at a time.

Procedure

The procedure for fitting multiple variables is as follows:

1. Open a spreadsheet with existing data for fitting.
2. Select the data you wish to fit (data should be in multiple columns with multiple rows).
3. Select *Simulation | Tools | Distributional Fitting (Multi-Variable)*.
4. Review the data, choose the types of distributions you want to fit to, and click OK.

Notes

Notice that the statistical ranking methods used in the distributional fitting routines are the chi-square test and Kolmogorov–Smirnov test. The former is used to test discrete distributions and the latter continuous distributions. Briefly, a hypothesis test coupled with the maximum likelihood procedure with an internal optimization routine is used to find the best-fitting parameters on each distribution tested and the results are ranked from the best fit to the worst fit. There are other distributional fitting tests such as the Anderson–Darling, Shapiro–Wilks, and others; however, these tests are very sensitive parametric tests and are highly inappropriate in Monte Carlo simulation distribution-fitting routines when different distributions are being tested. Due to their parametric requirements, these tests are most suited for testing normal distributions and distributions with normal-like behaviors (e.g., binomial distribution with a high number of trials and symmetrical probabilities) and will provide less accurate results when performed on nonnormal distributions. Take great care when using such parametric tests. The Kolmogorov–Smirnov and chi-square tests employed in Risk Simulator are nonparametric and semiparametric in nature, and are better suited for fitting normal and nonnormal distributions.

BOOTSTRAP SIMULATION

Theory

Bootstrap simulation is a simple technique that estimates the reliability or accuracy of forecast statistics or other sample raw data. Bootstrap simulation can be used to answer a lot of confidence and precision-based questions in simulation. For instance, suppose an identical model (with identical assumptions and forecasts but without any random seeds) is run by 100 different people. The results will clearly be slightly different. The question is, if we collected all the statistics from these 100 people, how will the mean be distributed, or the median, or the skewness, or excess kurtosis? Suppose one person has a mean value of, say, 1.50, while another 1.52. Are these two values statistically significantly different from one another or are they statistically similar and the slight difference is due entirely to random chance? What about 1.53? So, how far is far enough to say that the values are statistically different? In addition, if a model's resulting skewness is –0.19, is this forecast distribution negatively skewed or is it statistically close enough to zero to state that this distribution is symmetrical and not skewed? Thus, if we bootstrapped this forecast 100 times, that is, run a 1,000-trial simulation for 100 times and collect the 100 skewness coefficients, the skewness distribution would indicate how far zero is away from –0.19. If the 90 percent confidence on the bootstrapped skewness distribution contains the value zero, then we can state that on a 90 percent confidence level, this distribution is symmetrical and not skewed, and the value –0.19 is statistically close enough to zero. Otherwise, if zero falls outside of this 90 percent confidence area, then this distribution is negatively skewed. The same analysis can be applied to excess kurtosis and other statistics.

Essentially, bootstrap simulation is a hypothesis testing tool. Classical methods used in the past relied on mathematical formulas to describe the accuracy of sample statistics. These methods assume that the distribution of a sample statistic approaches a normal distribution, making the calculation of the statistic's standard error or confidence interval relatively easy. However, when a statistic's sampling distribution is not normally distributed or easily found, these classical methods are difficult to use. In contrast, bootstrapping analyzes sample statistics empirically by repeatedly sampling the data and creating distributions of the different statistics from each sampling. The classical methods of hypothesis testing are available in Risk Simulator and are explained in the next section. Classical methods provide higher power in their tests but rely on normality assumptions and can only be used to test the mean and variance of a distribution, as compared to bootstrap simulation, which provides lower power but is nonparametric and distribution-free, and can be used to test any distributional statistic.

Procedure

Use the following steps to run a bootstrap simulation:

1. Run a simulation with assumptions and forecasts.
2. Select *Simulation* | *Tools* | *Nonparametric Bootstrap*.
3. Select only *one* forecast to bootstrap, select the statistic(s) to bootstrap, and enter the number of bootstrap trials and click OK (Figure 6.15).

Results Interpretation

Figure 6.16 illustrates some sample bootstrap results. The example file used was *Hypothesis Testing and Bootstrap Simulation*. For instance, the 90 percent confidence for the skewness statistic is between –0.0189 and 0.0952, such that the value 0 falls within this confidence, indicating that on a 90 percent confidence, the skewness of this forecast is not statistically significantly different from zero, or that this distribution can be considered as symmetrical and not skewed. Conversely, if the value 0 falls outside of this confidence, then the opposite is true: The distribution is skewed (positively skewed if the forecast statistic is positive, and negatively skewed if the forecast statistic is negative).

Notes

The term *bootstrap* comes from the saying, "to pull oneself up by one's own bootstraps," and is applicable because this method uses the distribution of statistics themselves to analyze the statistics' accuracy. Nonparametric simulation is simply randomly picking golf balls from a large basket with

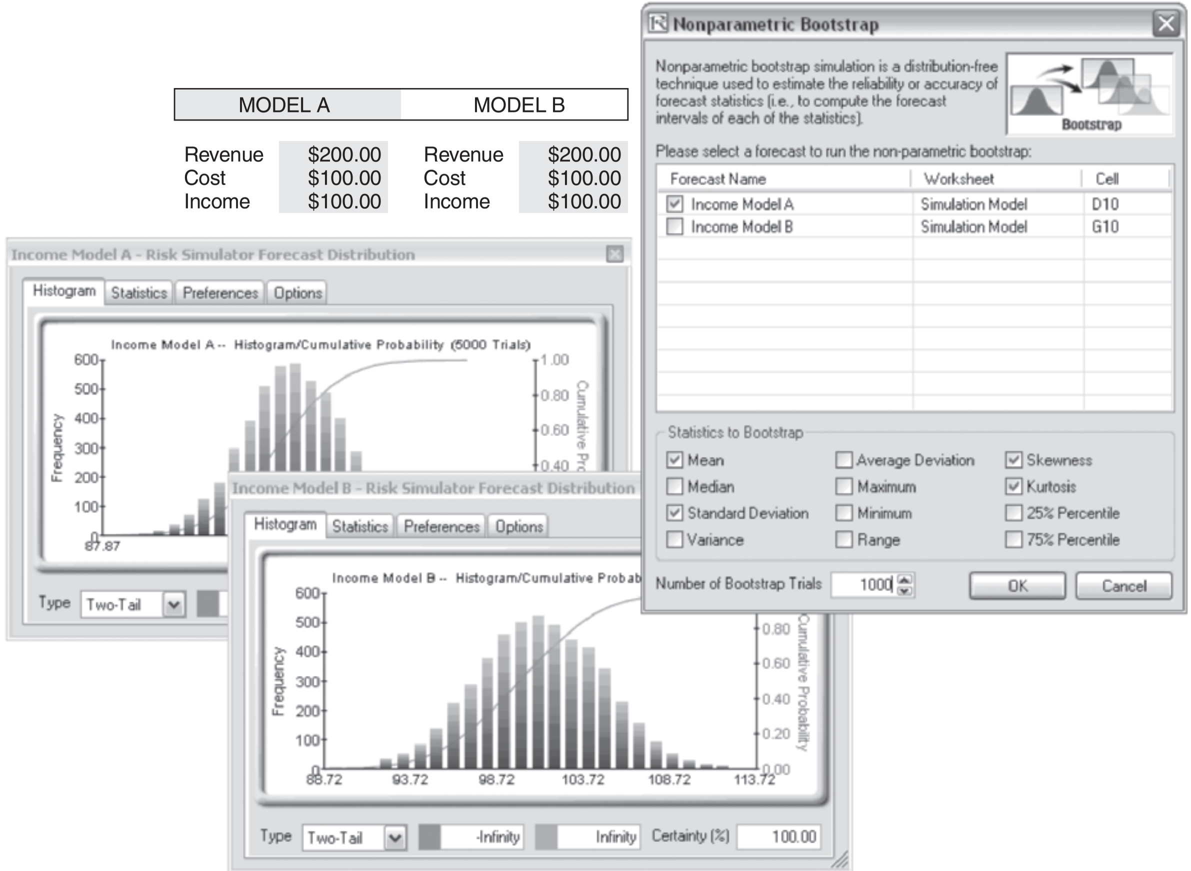


FIGURE 6.15 Nonparametric bootstrap simulation.

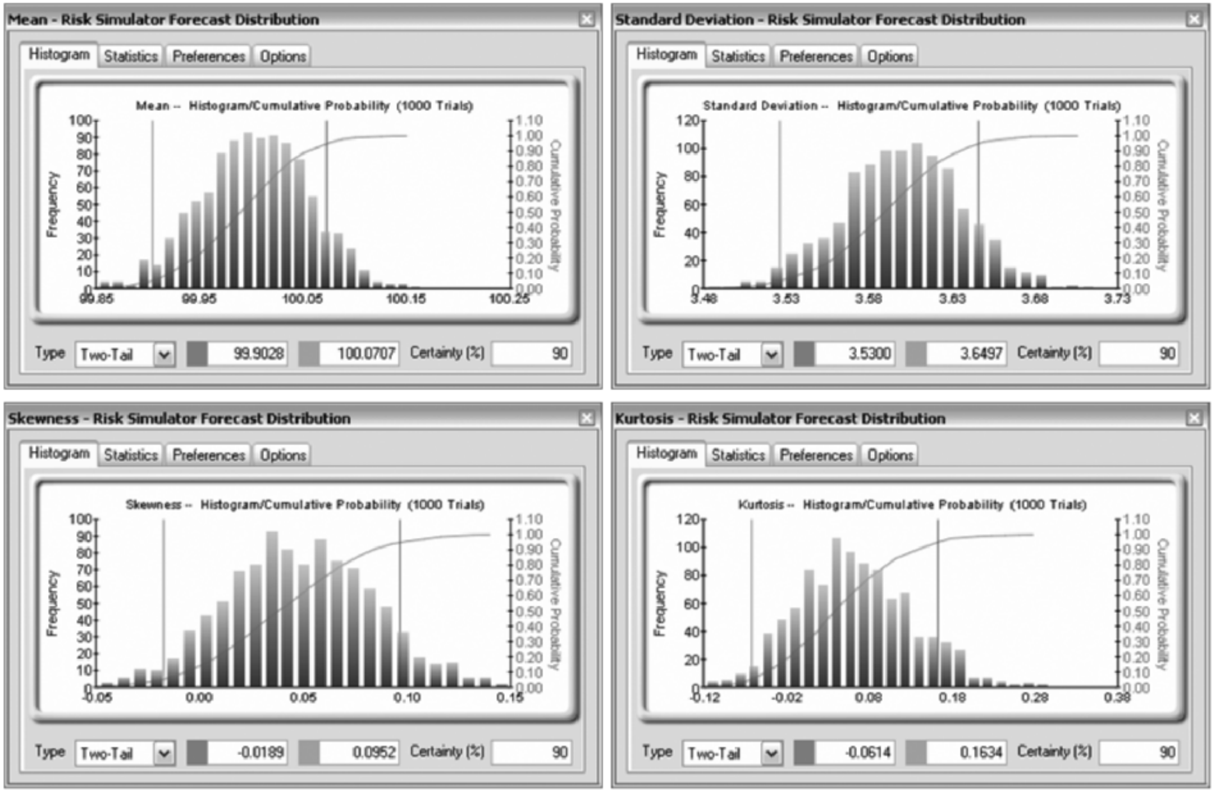


FIGURE 6.16 Bootstrap simulation results.

replacement where each golf ball is based on a historical data point. Suppose there are 365 golf balls in the basket (representing 365 historical data points). Imagine if you will that the value of each golf ball picked at random is written on a large whiteboard. The results of the 365 balls picked with replacement are written in the first column of the board with 365 rows of numbers. Relevant statistics (mean, median, mode, standard deviation, and so forth) are calculated on these 365 rows. The process is then repeated, say, five thousand times. The whiteboard will now be filled with 365 rows and 5,000 columns. Hence, 5,000 sets of statistics (that is, there will be 5,000 means, 5,000 medians, 5,000 modes, 5,000 standard deviations, and so forth) are tabulated and their distributions shown. The relevant *statistics of the statistics* are then tabulated, where from these results one can ascertain how confident the simulated statistics are. Finally, bootstrap results are important because according to the *Law of Large Numbers* and *Central Limit Theorem* in statistics, the mean of the sample means is an unbiased estimator and approaches the true population mean when the sample size increases.

HYPOTHESIS TESTING

Theory

A hypothesis test is performed when testing the means and variances of two distributions to determine if they are statistically identical or statistically different from one another; that is, to see if the differences between the means and variances of two different forecasts that occur are based on random chance or if they are, in fact, statistically significantly different from one another.

This analysis is related to bootstrap simulation with several differences. Classical hypothesis testing uses mathematical models and is based on theoretical distributions. This means that the precision and power of the test is higher than bootstrap simulation's empirically based method of simulating a simulation and letting the data tell the story. However, classical hypothesis test is only applicable for testing two distributions' means and variances (and by extension, standard deviations) to see if they are statistically identical or different. In contrast, nonparametric bootstrap simulation can be used to test for any distributional statistics, making it more useful, but the drawback is its lower testing power. Risk Simulator provides both techniques from which to choose.

Procedure

Use the following steps to run a hypothesis test:

1. Run a simulation with at least two forecasts.

2. Select *Simulation | Tools | Hypothesis Testing*.
3. Select the two forecasts to test, select the type of hypothesis test you wish to run, and click OK (Figure 6.17).

Results Interpretation

A two-tailed hypothesis test is performed on the null hypothesis (H_o) such that the two variables' population means are statistically identical to one another. The alternative hypothesis (H_a) is such that the population means are statistically different from one another. If the calculated p-values are less than or equal to 0.01, 0.05, or 0.10 alpha test levels, it means that the null hypothesis is rejected, which implies that the forecast means are statistically significantly different at the 1 percent, 5 percent, and 10 percent significance levels. If the null hypothesis is not rejected when the p-values are high, the means of the two forecast distributions are statistically similar to one another. The same analysis is performed on variances of two forecasts at a time using the pairwise F-test. If the p-values are small, then the variances (and standard deviations) are statistically different from one another. Otherwise, for large p-values, the variances are statistically identical to one another. See Figure 6.18. The example file used was *Hypothesis Testing and Bootstrap Simulation*.

Notes

The two-variable t-test with unequal variances (the population variance of forecast 1 is expected to be different from the population variance of forecast 2) is appropriate when the forecast distributions are from different populations (e.g., data collected from two different geographical locations or two different operating business units). The two-variable t-test with equal variances (the population variance of forecast 1 is expected to be equal to the population variance of forecast 2) is appropriate when the forecast distributions are from similar populations (e.g., data collected from two different engine designs with similar specifications). The paired dependent two-variable t-test is appropriate when the forecast distributions are from exactly the same population and subjects (e.g., data collected from the same group of patients before an experimental drug was used and after the drug was applied).

DATA EXTRACTION, SAVING SIMULATION RESULTS, AND GENERATING REPORTS

A simulation's raw data can be very easily extracted using Risk Simulator's *Data Extraction* routine. Both assumptions and forecasts can be extracted,

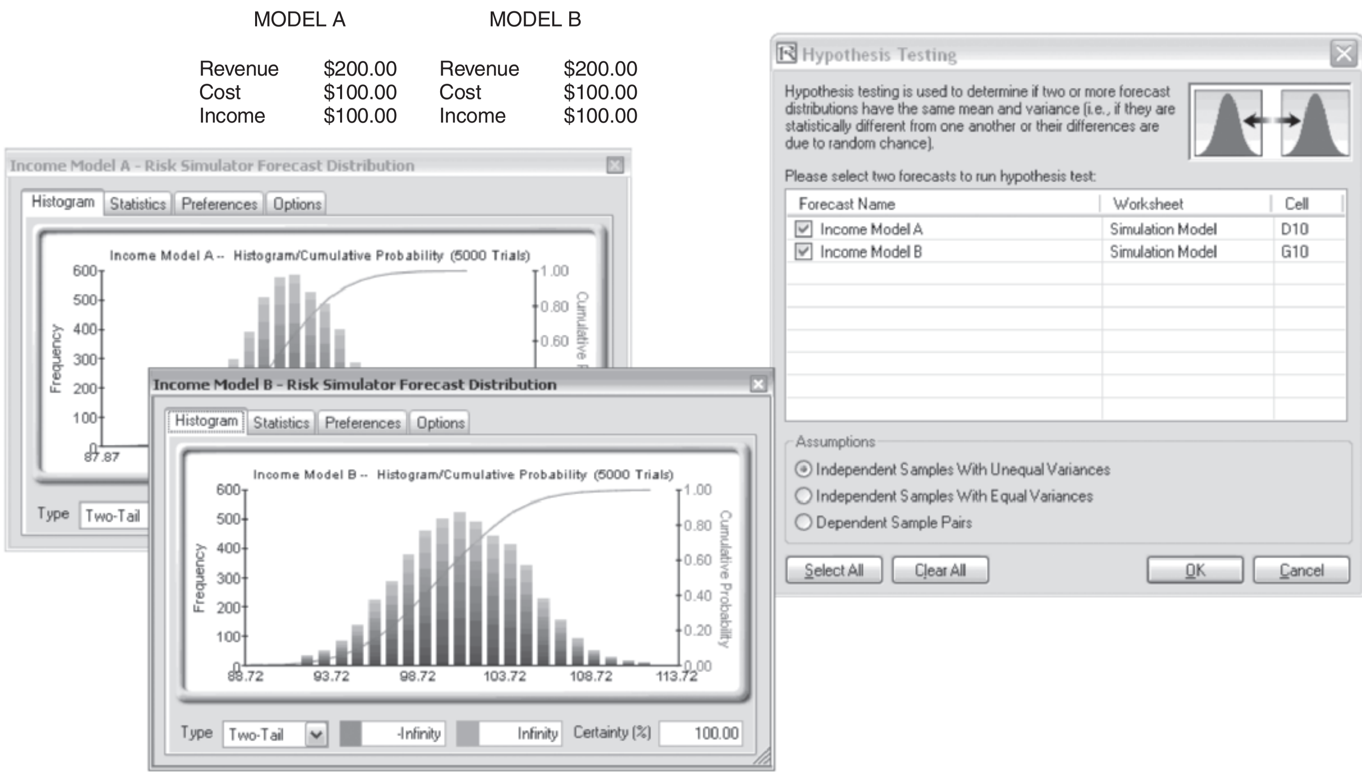


FIGURE 6.17 Hypothesis testing.

Hypothesis Test on the Means and Variances of Two Forecasts

Statistical Summary

A hypothesis test is performed when testing the means and variances of two distributions to determine if they are statistically identical or statistically different from one another; that is, to see if the differences between two means and two variances that occur are based on random chance or they are, in fact, different from one another. The two-variable t-test with unequal variances (the population variance of forecast 1 is expected to be different from the population variance of forecast 2) is appropriate when the forecast distributions are from different populations (e.g., data collected from two different geographical locations, two different operating business units, and so forth). The two variable t-test with equal variances (the population variance of forecast 1 is expected to be equal to the population variance of forecast 2) is appropriate when the forecast distributions are from similar populations (e.g., data collected from two different engine designs with similar specifications, and so forth). The paired dependent two-variable t-test is appropriate when the forecast distributions are from similar populations (e.g., data collected from the same group of customers but on different occasions, and so forth).

A two-tailed hypothesis test is performed on the null hypothesis H_o such that the two variables' population means are statistically identical to one another. The alternative hypothesis is that the population means are statistically different from one another. If the calculated p-values are less than or equal to 0.01, 0.05, or 0.10, the hypothesis is rejected, which implies that the forecast means are statistically significantly different at the 1%, 5%, and 10% significance levels. If the null hypothesis is not rejected when the p-values are high, the means of the two forecast distributions are statistically similar to one another. The same analysis is performed on variances of two forecasts at a time using the pairwise F-Test. If the p-values are small, then the variances (and standard deviations) are statistically different from one another, otherwise, for large p-values, the variances are statistically identical to one another.

Result

Hypothesis Test Assumption	Unequal Variances
Computed t-statistic:	–0.32947
P-value for t-statistic:	0.74181
Computed F-statistic:	1.026723
P-value for F-statistic:	0.351212

FIGURE 6.18 Hypothesis testing results.

but a simulation must first be run. The extracted data can then be used for a variety of other analyses and the data can be extracted to different formats—for use in spreadsheets, databases, and other software products.

Procedure

To extract a simulation's raw data, use the following steps:

1. Open or create a model, define assumptions and forecasts, and run the simulation.
2. Select *Simulation | Tools | Data Extraction*.
3. Select the assumptions and/or forecasts you wish to extract the data from and click OK.

The simulated data can be extracted and saved to an Excel worksheet, a text file (for easy import into other software applications), or as a *RiskSim* file, which can be reopened as Risk Simulator forecast charts at a later date. Finally, you can create a simulation report of all the assumptions and forecasts in your model by going to *Simulation | Create Report*. This is an efficient way to gather all the simulation inputs in one concise report.

CUSTOM MACROS

Simulation can also be run while harnessing the power of Visual Basic for Applications (VBA) in Excel. For instance, the examples in Chapter 2 on running models with VBA codes can be used in tandem with Risk Simulator. For an illustration of how to set the macros or customized functions to run with simulation, see the VBA Macro hands-on exercise (Retirement Funding with Inflation) at the end of this chapter.

APPENDIX—GOODNESS-OF-FIT TESTS

Several statistical tests exist for deciding if a sample set of data comes from a specific distribution. The most commonly used are the Kolmogorov–Smirnov test and the chi-square test. Each test has its advantages and disadvantages. The following sections detail the specifics of these tests as applied in distributional fitting in Monte Carlo simulation analysis. These two tests are used in Risk Simulator's distributional fitting routines.

Other goodness-of-fit tests such as the Anderson–Darling, Lilliefors, Jacque–Bera, Wilkes–Shapiro, and others are not used as these are parametric tests and their accuracy depends on the data set being normal or near-normal. Therefore, the results of these tests are oftentimes suspect or yield inconsistent results.

Kolmogorov–Smirnov Test

The Kolmogorov–Smirnov (KS) test is based on the empirical distribution function of a sample data set and belongs to a class of *nonparametric tests*. This nonparametric characteristic is the key to understanding the KS test, which simply means that the distribution of the KS test statistic does not depend on the underlying cumulative distribution function being tested. Nonparametric simply means no predefined distributional parameters are required. In other words, the KS test is applicable across a multitude of underlying distributions. Another advantage is that it is an exact test as compared to the chi-square test, which depends on an adequate sample size for the approximations to be valid. Despite these advantages, the KS test has several important limitations. It only applies to continuous distributions, and it tends to be more sensitive near the center of the distribution than at the distribution's tails. Also, the distribution must be fully specified.

Given N ordered data points $Y_1, Y_2, \ldots Y_N$, the empirical distribution function is defined as $E_n = n(i)/N$ where $n(i)$ is the number of points less than Y_i where Y_i values are ordered from the smallest to the largest value. This is a step function that increases by $1/N$ at the value of each ordered data point.

The null hypothesis is such that the data set follows a specified distribution while the alternate hypothesis is that the data set does not follow the specified distribution. The hypothesis is tested using the KS statistic defined as

$$KS = \max_{1 \le i \le N} \left| F(Y_i) - \frac{i}{N} \right|$$

where F is the theoretical cumulative distribution of the continuous distribution being tested that must be fully specified (i.e., the location, scale, and shape parameters cannot be estimated from the data).

As the null hypothesis is that the data follows some specified distribution, when applied to distributional fitting in Risk Simulator, a low p-value (e.g., less than 0.10, 0.05, or 0.01) indicates a bad fit (the null hypothesis is rejected) while a high p-value indicates a statistically good fit.

Chi-Square Test

The chi-square (CS) goodness-of-fit test is applied to binned data (i.e., data put into classes), and an attractive feature of the CS test is that it can be applied to any univariate distribution for which you can calculate the cumulative distribution function. However, the values of the CS test statistic are dependent on how the data is binned and the test requires a sufficient sample size in order for the CS approximation to be valid. This test is sensitive to the choice of bins. The test can be applied to discrete distributions such as the binomial and the Poisson, while the KS test is restricted to continuous distributions.

The null hypothesis is such that the data set follows a specified distribution while the alternate hypothesis is that the data set does not follow the specified distribution. The hypothesis is tested using the CS statistic defined as

$$\chi^2 = \sum_{i=1}^{k} (O_i - E_i)^2 \; / \; E_i$$

where O_i is the observed frequency for bin i and E_i is the expected frequency for bin i. The expected frequency is calculated by

$$E_i = N(F(Y_U) - F(Y_L))$$

where F is the cumulative distribution function for the distribution being tested, Y_U is the upper limit for class i, Y_L is the lower limit for class i, and N is the sample size.

The test statistic follows a CS distribution with $(k - c)$ degrees of freedom where k is the number of nonempty cells and c = the number of esti-

TABLE 6.2 Chi-Square Test

Alpha Level (%)	Cutoff
10	32.00690
5	35.17246
1	41.63840

Note: Chi-square goodness-of-fit test sample critical values. Degrees of freedom 23.

mated parameters (including location and scale parameters and shape parameters) for the distribution + 1. For example, for a three-parameter Weibull distribution, $c = 4$. Therefore, the hypothesis that the data are from a population with the specified distribution is rejected if $\chi^2 > \chi^2(\alpha, k - c)$ where $\chi^2(\alpha, k - c)$ is the CS percent point function with $k - c$ degrees of freedom and a significance level of α (see Table 6.2).

Again, as the null hypothesis is such that the data follows some specified distribution, when applied to distributional fitting in Risk Simulator, a low p-value (e.g., less than 0.10, 0.05, or 0.01) indicates a bad fit (the null hypothesis is rejected) while a high p-value indicates a statistically good fit.

QUESTIONS

1. Name the key similarities and differences between a tornado chart and a spider chart. Then, compare tornado and spider charts with sensitivity analysis.
2. In distributional fitting, sometimes you may not get the distribution you thought is the right fit as the best choice. Why is this so? Also, why does the beta distribution usually come up as one of the top few candidates as the best-fitting distribution?
3. Briefly explain what a hypothesis test is.
4. How is bootstrap simulation related to precision and error control in simulation?
5. In sensitivity analysis, how is percent variation explained linked to rank correlation?

Additional hands-on exercises are presented in the following pages. These exercises require Risk Simulator to be installed and application of the techniques presented in this chapter.

Real Options Valuation

Risk Simulator
Hands-On Exercise

Tornado, Spider and Sensitivity Charts (Nonlinear)

This sample model is used to illustrate how to use Risk Simulator for:

1. Running a pre-simulation sensitivity analysis (Tornado and Spider Charts)
2. Running a post-simulation sensitivity analysis (Sensitivity Charts)

Model Background

File Name: Tornado and Sensitivity Charts (Nonlinear).xls

This example illustrates a simple option valuation model and shows how sensitivity analysis can be performed prior to running a simulation and after a simulation is run. Tornado and Spider charts are static sensitivity analysis tools useful for determining which variables impact the key results the most, i.e., each precedent variable is perturbed a set amount and the key result is analyzed to determine which input variables are the critical success factors with the most impact. Whereas sensitity charts are dynamic, in that all precedent variables are perturbed together in a simultaneous fashion (the effects of autocorrelations, cross-correlations, and interactions are all captured in the resulting sensitivity chart). The model is summarized below:

Generalized Black-Scholes Model

Stock Price	$100.00
Strike Price	$100.00
Maturity in Years	5.00
Risk-free Rate	5.00%
Annualized Volatility	25.00%
Dividend Yield	2.00%
Black-Scholes Result	***$25.48***

$$Call = Se^{-qT}\Phi\left[\frac{\ln(S/X)+(rf-q+\sigma^2/2)T}{\sigma\sqrt{T}}\right]-Xe^{-rf(T)}\Phi\left[\frac{\ln(S/X)+(rf-q-\sigma^2/2)T}{\sigma\sqrt{T}}\right]$$

Creating a Tornado and Sensitivity Chart

To run this model, simply:

1. Go to the *Black-Scholes Model* worksheet and select the *Black-Scholes* result (cell *E13*)
2. Select **Simulation | Tools | Tornado Analysis** (or click on the **Tornado Chart icon**)
3. Check that the software's intellegent naming is correct for the precedent values
4. Change the upper and lower test values from 10% to 50% on all variables and click **OK**

Interpreting the Results

The report generated illustrates the sensitivity table (starting base value of the key variable as well as the perturbed values and the precedents). The precedent with the highest impact (range of output) is listed first. The Tornado chart illustrates this analysis graphically. The Spider chart performs the same analysis but also accounts for nonlinear effects, i.e., if the input variables have a nonlinear effect on the output variable, the lines on the Spider chart will be curved. Notice that the stock price and strike price have the highest effects (positive and negative effects respectively) and the inputs are nonlinear. Positively-sloped lines indicate positive relationships, while negatively-sloped lines indicate negative relationships. The steeper the slope of the line, the higher the impact on the bottom line value.

Creating a Sensitivity Chart

To run this model, simply:

1. Open the existing simulation profile called *Sensitivity* (select **Simulation | Change Profile | Sensitivity**)
2. Run the simulation (**Simulation | Run Simulation**)
3. Select **Simulation | Tools | Sensitivity Analysis**

Interpreting the Results

Notice that if correlations are turned off, the results of the Sensitivity chart is similar to the Tornado chart. Now, reset the simulation, and turn on correlations (select **Simulation | Reset Simulation**, then **Simulation | Edit Profile** and check *Apply Correlations*, then **Simulation | Run Simulation**), and repeat the steps above for creating a Sensitivity chart. Notice that when correlations are applied, the resulting analysis may be slightly different due to the interactions among variables. Notice that when correlations are turned on, the Stock Price position is reduced to a less prominent effect as it is correlated to Dividend Yield, a less dominant variable. Note that in the Sensitivity chart, Volatility has a larger effect once the interactions among variables have all been accounted for. Volatility in real life is a key indicator of option value and the analysis here proves this fact.

Data Fitting

This sample model is used to illustrate how to use Risk Simulator for:

1. Fitting a single variable to existing data
2. Fitting multiple variables to existing data
3. Simulating, extracting data, and refitting to distributions

Model Background
File Name: Data Fitting.xls
This example is used to illustrate how existing sample data can be used to find the statistically best-fitting distribution. By doing so, we also confirm the simulation results through the distributional fitting routine, that is, we simulate a particular distribution, extract its raw data, and refit it back to all distributions. The sample model looks like the following:

Sample Simulation for Continuous Variable Fit:		**Output**	**Distributional Assumptions:**
Normal	100.00	100.00	Mean = 100, Stdev = 10
Uniform	50.00	50.00	Min = 45, Max = 55
Sample Simulation for Discrete Distribution Fit:			
Binomial	10	10	N = 10, P = 0.75

Random Seed is set at 123456 for 200 trials

A 200-trial simulation is then run and the resulting values are extracted. A sample of the data is seen on the right:

Normal	Uniform	Binomial
87.53	53.30	8.00
97.25	45.96	7.00
87.41	52.12	8.00
85.99	51.74	9.00
95.55	51.98	8.00
102.69	53.39	6.00
99.98	46.76	8.00
83.39	52.02	8.00
95.41	46.97	9.00
88.76	47.75	8.00
91.06	49.15	8.00
99.58	46.60	6.00

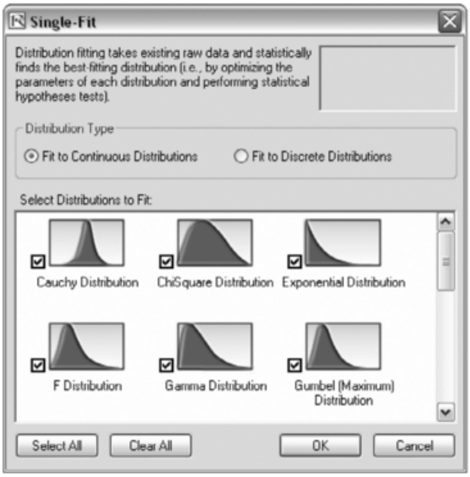
Then, using the 200 data points for each variable, we perform a distributional fit using Risk Simulator's Single-Fit:

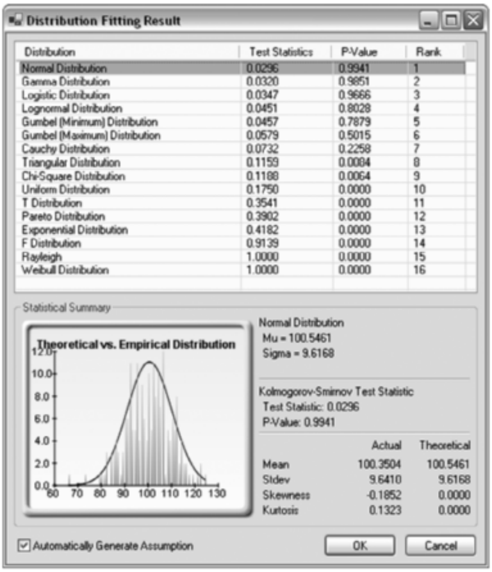
Running a Single-Fit

To run this model, simply:

1. Go to the *Raw Data worksheet* and *select cells C2:C201*.
2. Click on **Simulation | Tools | Distributional Fitting (Single Variable)**.
3. Make sure *Fit Continuous Distributions* is selected and all distributions are checked, then click **OK**.
4. The resulting fit of all distributions are shown. Select the best fit (ranked first), view the statistics, and click **OK**.
5. A report will be generated indicating all the relevant statistics as well as the data used for fitting (for future reference).

Note that if a *profile* exists and if the *Automatically Generate Assumption* choice is selected, then the report will contain an assumption that is the best fit. Otherwise, only the type of distribution and its relevant input assumptions are provided. You can repeat this exercise on the remaining data points provided.





Notes:

You can replicate the fitting routine on multiple variables simultaneously (variables must be arranged in columns) instead of one at a time. Select the area C2:I201 and run the multiple variable fitting routine by selecting **Simulation | Tools | Distributional Fitting (Multi-Variable)**.

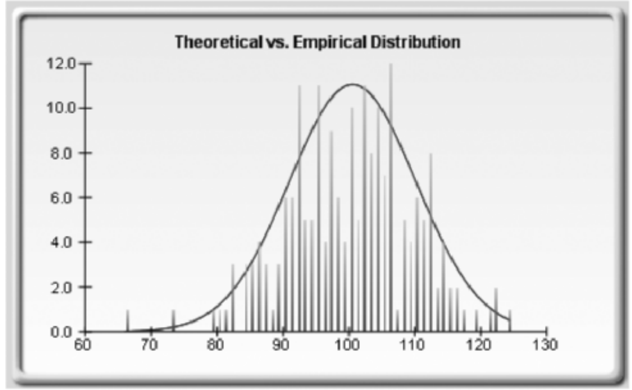
You can replicate the data set by opening the existing simulation profile (**Simulation | Change Profile**), run a simulation (**Simulation | Run**), then extract the raw data (**Simulation | Tools | Data Extraction**). Then, you can try fitting the raw data to the distributions.

However, there are several key points to remember. First, more data implies a better statistical fit. Do not fit to very few data points and expect a good fit. Second, only positive discrete data (integers) can be fitted to discrete distributions. When negative values or continuous data exist, you should always fit to continuous distributions. Third, certain distributions are related to other distributions through their statistical properties. For example, a t distribution becomes a normal distribution when degrees of freedom is high, a poisson distribution can be used to approximate a binomial, or a normal can be used to approximate a poisson, hypergeometric, and binomial. There are many other such relationships and just because the fit is not exactly to the distribution expected does not mean the data is bad or the routine is incorrect. It simply means that another distribution is better suited for the data.

Single Variable Distributional Fitting

Statistical Summary

Fitted Assumption		100.54
Fitted Distribution		**Normal Distribution**
Mu		100.54
Sigma		9.62
Kolmogorov-Smirnov Statistic		0.03
P-Value for Test Statistic		0.9941
	Actual	Theoretical
Mean	100.35	100.55
Standard Deviation	9.64	9.62
Skewness	-0.19	0.00
Excess Kurtosis	0.13	0.00



Original Fitted Data

93.75	91.36	102.29	104.64	86.79	97.32	105.90	84.60	79.74	106.69	97.70	98.40	92.06
100.64	87.45	104.38	82.75	92.96	112.42	111.54	86.04	122.86	124.39	92.00	103.90	108.09
86.91	97.30	112.19	102.39	103.72	109.18	87.03	84.43	112.85	92.44	103.32	85.93	97.48
80.14	96.47	111.17	97.23	97.86	114.91	93.91	112.29	107.04	115.66	95.46	96.43	93.86
92.36	102.84	100.08	104.72	100.36	90.79	97.00	101.03	104.57	103.96	111.76	99.36	114.14
106.13	82.08	102.52	98.24	100.18	94.15	105.24	95.78	108.79	112.07	96.19	91.25	104.12
93.48	93.88	103.66	95.55	102.55	106.54	95.71	89.61	113.70	104.68	109.22	97.08	104.46
113.94	98.17	94.42	100.44	92.49	103.81	95.82	119.24	92.55	101.45	110.70	115.58	101.99
105.52	94.05	102.60	86.76	103.99	106.11	94.95	89.71	111.44	95.43	100.24	106.83	108.34
95.93	99.03	91.77	101.73	97.66	105.66	92.21	117.60	114.55	110.25	112.52	110.16	106.33
108.20	84.79	73.53	112.26	95.04	99.10	116.40	106.17	90.03	94.07	106.03	104.01	98.07
95.01	104.31	100.00	91.38	95.61	99.77	102.17	114.89	100.04	91.73	92.99	102.48	105.15
110.74	88.94	101.33	85.81	87.65	105.86	121.67	106.95	90.65	90.96	109.69	91.35	122.36
116.34	92.30	102.95	89.77	90.94	106.65	82.94	99.00	96.86	95.72	102.43	112.01	111.37
106.98	66.91	110.22	108.33	85.59	100.88	109.23	105.69	106.00	81.05	104.42	102.03	92.70
98.57	90.75	103.05	100.40	110.74								

Hypothesis Testing and Bootstrapping

This sample model is used to illustrate how to use Risk Simulator for:

1. Running a hypothesis test after a simulation
2. Running a hypothesis test with raw data
3. Understanding the concept of random seeds
4. Nonparametric bootstrap simulation

Model Background

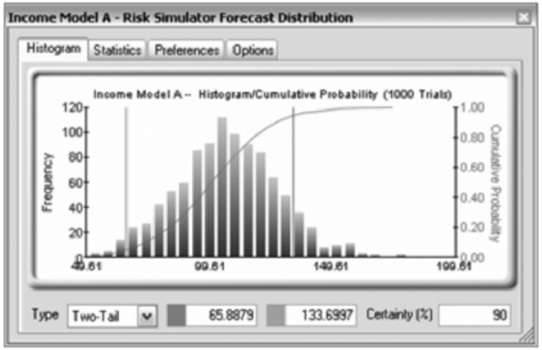
File Name: Hypothesis Testing and Bootstrap Simulation.xls

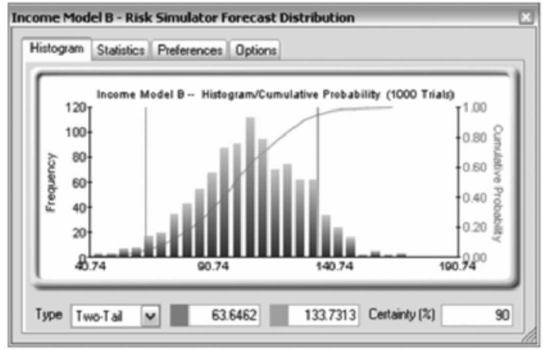
This example is used to show how two different forecasts or sets of data can be tested against one another to determine if they have have the same means and variances. That is, if the first distribution has a mean of 100, how far away does the mean of the second distribution have to be such that they are considered statistically different? The following illustrates two models (A and B) with the same calculations (see Simulation Model worksheet) where the income is revenue minus cost. Both sets of models have the same inputs and the same distributional assumptions on the inputs, and the simulation is run on the random seed of 123456. There are two major items that are noteworthy. The first is that the means and variances (as well as standard deviations) are slightly different. This raises the question as to whether the means and variances of these two distributions are identical.

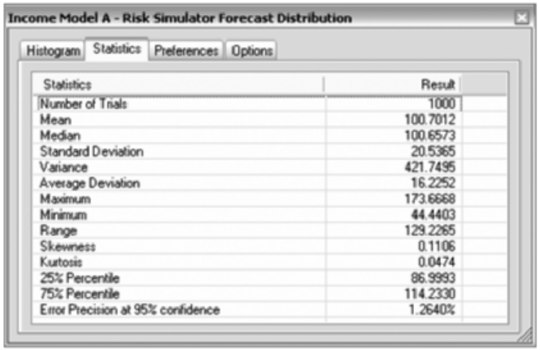
A hypothesis test can be applied to answer this first question. A nonparametric bootstrap simulation can also be applied to test the other statistics to see if they are statistically valid. The second item of interest is that the results from A and B are different although the input assumptions are identical and an overall random seed has been applied. This is because with a random seed applied, each distribution is allowed to vary independently as long as it is not correlated to another variable. This is a key and useful fact in Monte Carlo simulation.

The model is seen on the right:

	MODEL A		MODEL B
Revenue	$ 200.00	Revenue	$ 200.00
Cost	$ 100.00	Cost	$ 100.00
Income	$ 100.00	Income	$ 100.00

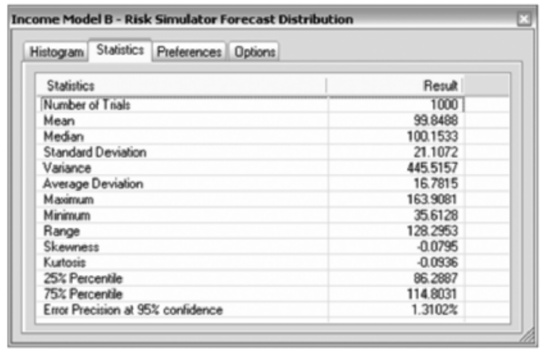





Income Model A - Risk Simulator Forecast Distribution

Histogram | Statistics | Preferences | Options

Statistics	Result
Number of Trials	1000
Mean	100.7012
Median	100.6573
Standard Deviation	20.5365
Variance	421.7495
Average Deviation	16.2252
Maximum	173.6668
Minimum	44.4403
Range	129.2265
Skewness	0.1106
Kurtosis	0.0474
25% Percentile	86.9993
75% Percentile	114.2330
Error Precision at 95% confidence	1.2640%


Income Model B - Risk Simulator Forecast Distribution

Histogram | Statistics | Preferences | Options

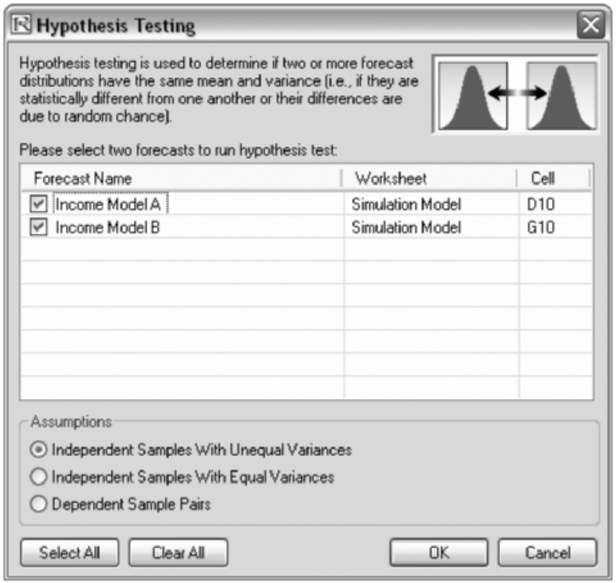
Statistics	Result
Number of Trials	1000
Mean	99.8488
Median	100.1533
Standard Deviation	21.1072
Variance	445.5157
Average Deviation	16.7815
Maximum	163.9081
Minimum	35.6128
Range	128.2953
Skewness	-0.0795
Kurtosis	-0.0936
25% Percentile	86.2887
75% Percentile	114.8031
Error Precision at 95% confidence	1.3102%

Running a Hypothesis Test

To run this model, simply:

1. First open the existing simulation profile by selecting **Simulation | Change Simulation Profile**. Choose the *Hypothesis Testing* profile.
2. Select **Simulation | Run Simulation** or click on the **Run icon**.
3. After the simulation run is complete, select **Simulation | Tools | Hypothesis Test**.
4. Make sure *both forecasts are selected* and click **OK**.

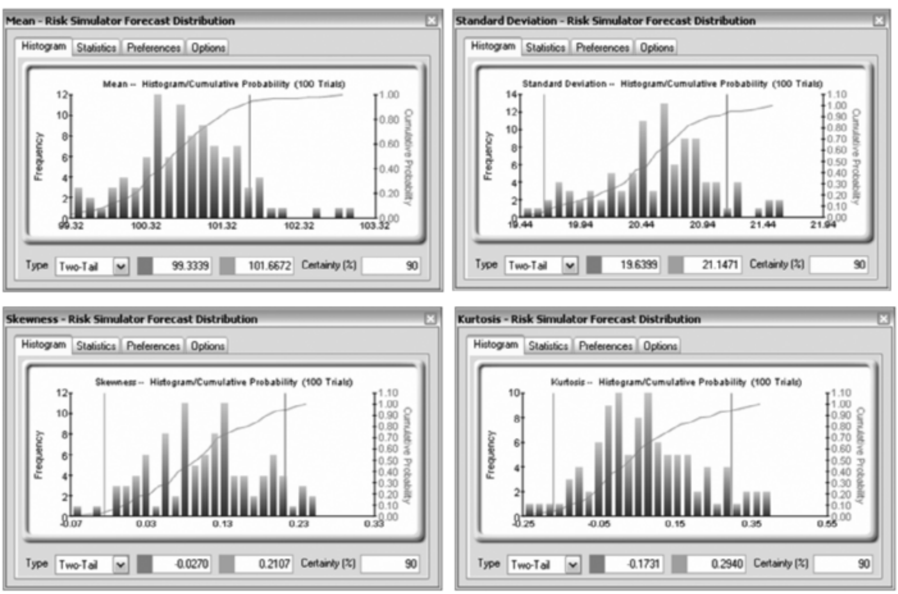
The report and results are provided in this example. The results indicate that the p-value for the t-test is higher that 0.05, indicating that both means are statistically identical, and that any variations are due to random white noise. Further, the p-value is also high for the F-test, indicating that both variances are also statistically identical to one another.



Running a Nonparametric Bootstrap Simulation

The hypothesis test above is a theoretical test and is hence more accurate. However, these tests do not exist for other statistics, hence, an empirical approach is required, namely, nonparametric bootstrap simulation. To run the bootstrap simulation, simply reset and rerun the simulation, then, once the simulation is complete, click on Simulation, Tools, Nonparametric Bootstrap. Choose one of the forecasts (only one forecast can be chosen at a time when running bootstrap simulation) and select the statistics of interest and click OK.

The resulting forecast charts are empirical distributions of the statistics. By typing in 90 on the certainty box and hitting tab on the keyboard, the 90 percent confidence is displayed for each statistic. For instance, the skewness interval is between -0.0270 and 0.2107, indicating that the value zero is within this interval, that is, at the 90 percent two-tail confidence (or significance of 0.10 two-tailed), model A does not have a skew.



Retirement Funding with Inflation

This sample model is used to illustrate how to use Risk Simulator for:

1. Running Monte Carlo simulation and capturing the output results
2. Running VBA macros in your models together with Monte Carlo simulation

Model Background

File Name: Retirement Funding with VBA Macros.xls

This model illustrates a simplified retirement amortization computation allowing for inflation adjustments. That is, given how much you currently have and the expected rate of return on your investment vehicles, the expected inflation rates, your current salary, the expected amount you wish to have starting at retirement, this model computes how much annual investment is required to sufficiently fund your retirement. A sample snapshot of the model follows:

RETIREMENT FUNDING (INFLATION ADJUSTED)

Input Data

Initial Savings Amount	$100,000.00	Annual Salary	$120,000.00
Annual Required Rate of Return	8.00%	Replacement Ratio at Retirement	70%
Current Age	32.00	*(The percentage of your current income required at retirement)*	
Retirement Age	55.00	Inflation Rate	2%
Average Terminal Age	90.00	Terminal Wealth	($0)
		Terminal Wealth Goal	$0.00
		Annual Contribution	$11,180.87

PAYOUT SCHEDULE

Age	Status	Starting Value	Contributions	Withdrawal	Wealth
32	Working	$100,000.00	$11,180.87	$0.00	$111,180.87
33	Working	$108,000.00	$23,256.20	$0.00	$131,256.20
34	Working	$116,640.00	$36,297.56	$0.00	$152,937.56
35	Working	$125,971.20	$50,382.23	$0.00	$176,353.43
36	Working	$136,048.90	$65,593.67	$0.00	$201,642.57
37	Working	$146,932.81	$82,022.03	$0.00	$228,954.84
38	Working	$158,687.43	$99,764.66	$0.00	$258,452.09
39	Working	$171,382.43	$118,926.70	$0.00	$290,309.12
40	Working	$185,093.02	$139,621.70	$0.00	$324,714.72

Annual Required Rate of Return	This is the expected rate of return on the investment vehicle
Current Age	Your current age right now, corresponding to the initial savings
Retirement Age	Your proposed age at retirement
Average Terminal Age	The expected age at death
Annual Salary	Your current salary level
Replacement Ratio	Percentage of your current salary required at retirement
Inflation Rate	Annualized inflation rate expected
Terminal Wealth Goal	The typical target is $0 (complete amortization)

Computed Results:

Terminal Wealth	How much you have left at terminal age
Annual Contribution	Required annual contribution to have sufficient funds for retirement

Monte Carlo Simulation

As rate of returns on investment and inflation rate are both uncertain and varies, these two variables are defined as input assumptions. The output forecast is the required annual contribution. This model already has a simulation profile, assumptions, forecasts, and VBA macros created.

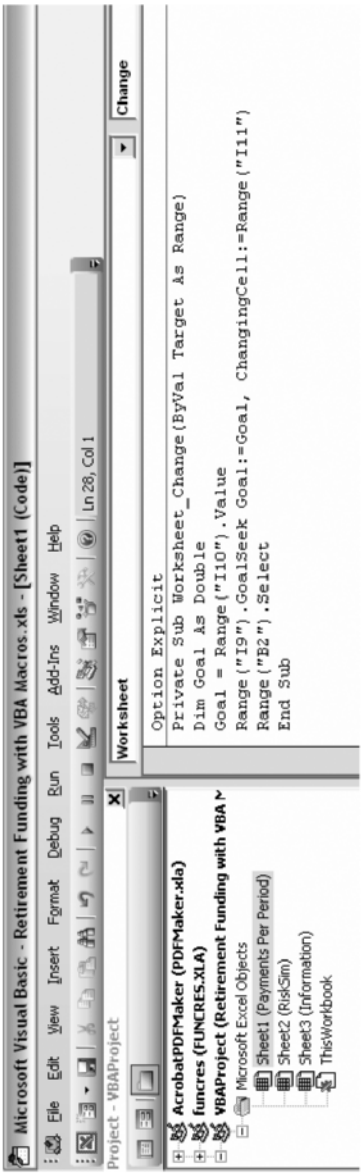
To run this model, simply:

1. Go to the *Payments Per Period* worksheet
2. Select **Simulation** | **Change Simulation Profile** and select *Risk Simulator* profile
3. Click on the **RUN** icon or **Simulation** | **Run Simulation**

Model Analysis

You can recreate the model's simulation and VBA analysis by repeating the following steps:

1. Click on Alt-F11 to get to the Excel VBA environment.
2. Double click on Sheet 1 and you will see the VBA macro code used to compute the Annual Contribution. Notice that the codes are listed under the Private Sub called Worksheet Change. This means that every time a value changes in the worksheet (i.e., when a simulation trial occurs), the macro will be executed and the value of the annual contribution will be computed accordingly in each simulation trial.
3. Exit the VBA environment and go back to the Payments Per Period worksheet.
4. Create a New Simulation Profile and provide it an appropriate name.
5. Define Assumptions on the Interest Annual Required Rate of Return and Inflation Rate.
6. Define Forecast on Annual Contribution.
7. Run the simulation.



Microsoft Visual Basic - Retirement Funding with VBA Macros.xls - [Sheet1 (Code)]
File Edit View Insert Format Debug Run Tools Add-Ins Window Help
Ln 28, Col 1
Project - VBAProject
AcrobatPDFMaker (PDFMaker.xla)
funcres (FUNCRES.XLA)
VBAProject (Retirement Funding with VBA M
Microsoft Excel Objects
Sheet1 (Payments Per Period)
Sheet2 (RiskSim)
Sheet3 (Information)
ThisWorkbook
Worksheet
Change
Option Explicit
Private Sub Worksheet_Change(ByVal Target As Range)
Dim Goal As Double
Goal = Range("I10").Value
Range("I9").GoalSeek Goal:=Goal, ChangingCell:=Range("I11")
Range("B2").Select
End Sub

PART Four

Industry Applications

CHAPTER 7

Extended Business Cases I: Pharmaceutical and Biotech Negotiations, Oil and Gas Exploration, Financial Planning with Simulation, Hospital Risk Management, and Risk-Based Executive Compensation Valuation

This chapter provides the first installment of five extended business cases. The first case pertains to the application of Monte Carlo simulation and risk analysis in the biotech and pharmaceutical industries. The case details the use of risk analysis for deal making and structuring, and is contributed by Dr. Charles Hardy. The second case in this chapter is contributed by Steve Hoye, a veteran of the oil and gas industry. Steve details the risks involved in oil exploration and production by illustrating a comprehensive oil exploration case from cradle to grave. Then, a financial planning case is presented by Tony Jurado, in considering the risks involved in retirement planning. The next case illustrates how Monte Carlo simulation coupled with queuing theory can be applied to hospital planning, and is contributed by Larry Pixley, an expert consultant in the health-care sector. Finally, Patrick Haggerty illustrates how simulation can be used to engineer a risk-based executive compensation plan.

CASE STUDY: PHARMACEUTICAL AND BIOTECH DEAL STRUCTURING

This business case is contributed by Dr. Charles Hardy, principal of BioAxia Incorporated of Foster City, California, a consulting firm specializing in valuation and quantitative deal structuring for bioscience firms. He is also chief financial officer and director of business development at Panorama Research, a biotechnology incubator in the San Francisco Bay Area. Dr. Hardy has a Ph.D. in pathobiology from the University of Washington in Seattle, Washington, and an MBA in finance and entrepreneurship from the University of Iowa in Iowa City, Iowa. He has functioned in a variety of roles for several start-up companies, including being CEO of Pulmogen, an early-stage medical device company. Dr. Hardy lives and works in the San Francisco Bay Area.

Smaller companies in the biotechnology industry rely heavily on alliances with pharmaceutical and larger companies to finance their R&D expenditures. Pharmaceutical and larger organizations in turn depend on these alliances to supplement their internal R&D programs. In order for smaller organizations to realize the cash flows associated with these alliances, they must have a competent and experienced business development component to negotiate and structure these crucial deals. In fact, the importance of these business collaborations to the survival of most young companies is so great that deal-making experience, polished business-development skills, and a substantial network of contacts are all frequent assets of the most successful executives of start-up and early-stage biotechnology companies.

Although deal-making opportunities for biotech companies are abundant because of the pharmaceutical industry's need to keep a healthy pipeline of new products in development, in recent years deal-making opportunities have lessened. Intuitively, then, firms have to be much more careful in the way they structure and value the deals in which they do get the opportunity to participate. However, despite this importance, a large number of executives prefer to go with comparable business deal structures for these collaborations in the hope of maximizing shareholder value for their firms, or by developing deal terms using their own intuition rather than developing a quantitative methodology for deal valuation and optimization to supplement their negotiation skills and strategies. For companies doing only one deal or less a year, perhaps the risk might be lower by structuring a collaboration based on comparable deal structures; at least they will get as much as the average company, or will they?

As described in this case study, *Monte Carlo simulation*, *stochastic optimization*, and *real options* are ideal tools for valuing and optimizing the financial terms of collaborative biomedical business deals focused on the development of human therapeutics. A large amount of data associated with clinical trial stage lengths and completion probabilities are publicly available. By quantitatively valuing and structuring deals, companies of all sizes can gain maximum shareholder value at all stages of development, and, most importantly, future cash flows can be defined based on expected cashflow needs and risk preference.

Deal Types

Most deals between two biotechnology companies or a biotechnology company and pharmaceutical company are strategic alliances where a cooperative agreement is made between two organizations to work together in defined ways with the goal of successfully developing or commercializing products. As the following list describes, there are several different types of strategic alliances:

- *Product Licensing.* A highly flexible and widely applicable arrangement where one party wishes to access the technology of another organization with no other close cooperation. This type of alliance carries very low risk and these types of agreements are made at nearly every stage of pharmaceutical development.
- *Product Acquisition.* A company purchases an existing product license from another company and thus obtains the right to market a fully or partially developed product.
- *Product Fostering.* A short-term exclusive license for a technology or product in a specific market that will typically include hand-back provisions.
- *Comarketing.* Two companies market the same product under different trade names.
- *Copromotion.* Two parties promote the same product under the same brand name.
- *Minority Investment Alliance.* One company buys stock in another as part of a mutually desired strategic relationship.

The historical agreement valued and optimized in this case study is an example of a product-licensing deal.

Financial Terms

Each business deal is decidedly unique, which explains why no "generic" financial model is sufficient to value and optimize every opportunity and collaboration. A biomedical collaborative agreement is the culmination of the combined goals, desires, requirements, and pressures from both sides of the bargaining table, possibly biased in favor of one party by exceptional negotiating skills, good preparation, more thorough due diligence, and accurate assumptions, and less of a need for immediate cash.

The financial terms agreed on for licensing or acquiring a new product or technology depend on a variety of factors, most of which impact the value of the deal. These include but are not limited to:

- Strength of the intellectual property position.
- Exclusivity of the rights agreed on.
- Territorial exclusivity granted.
- Uniqueness of the technology transferred.
- Competitive position of the company.
- Stage of technology developed.
- Risk of the project being licensed or sold.

Although every deal is different, most include: (1) licensing and R&D fees; (2) milestone payments; (3) product royalty payments; and (4) equity investments.

Primary Financial Models

All calculations described in this case study are based on discounted cash-flow (DCF) principals using risk-adjusted discount rates. Here, assets under uncertainty are valued using the following basic financial equation:

$$NPV = \sum_{i=0}^{n} \frac{E(CF_t)}{(1 + r_t + \pi_t)^t}$$

where NPV is the net present value, $E(CF_t)$ is the expected value of the cash flow at time t, r_t is the risk-free rate, and π_t is the risk premium appropriate for the risk of CF_t.

All subcomponents of models described here use different discount rates if they are subject to different risks. In the case of biomedical collaborative agreements, all major subcomponents (licensing fees, R&D costs and funding, clinical costs, milestone payments, and royalties) are frequently subject to many different distinct risks, and thus are all assigned their own discount rates based on a combination of factors, with the subject company's weighted average cost of capital (WACC) used as the base value. To incorporate the uncertain and dynamic nature of these risk assumptions into the model, all of these discount rates are themselves Monte Carlo variables. This discounting supplementation is critical to valuing the deal accurately, and most important for later stochastic optimization.

Historical Deal Background and Negotiated Deal Structure

The deal valued and optimized in this case study was a preclinical, exclusive product-licensing agreement between a small biotechnology company and a larger organization. The biopharmaceutical being valued had one major therapeutic indication, with an estimated market size of $1 billion at the date the deal was signed. The licensee negotiated the right to sublicense. The deal had a variety of funding provisions, with a summary of the financial terms presented in Table 7.1. The licensor estimated they were approximately 2 years away from filing an investigational new drug (IND) application that would initiate clinical trials in humans. For the purposes of the deal valuation and optimization described here, it is assumed that no information asymmetries exist between the companies forming the collaboration (i.e., both groups feel there is an equally strong likelihood their candidate biopharmaceutical will be a commercial success).

Licensing fees for the historical deal consisted of an up-front fee followed by licensing maintenance fees including multipliers (Table 7.1). Licensing maintenance fees will terminate on any one of the following events: (1) first IND filing by licensor; (2) tenth anniversary of the effective date; and (3) termination of the agreement. Milestone values for the historical deal numbered only three, with a $500,000 payment awarded on IND filing, a $1,500,000 payment on new drug application (NDA) filing, and a $4,000,000 payment on NDA approval (Table 7.1). The negotiated royalties for the historical deal were a flat 2.0 percent of net sales.

As described later in this case, two additional deal scenarios were constructed and stochastically optimized from the historical structure: a higher-value, lower-risk (HVLR) scenario and a higher-value, higher-risk (HVHR) scenario (Table 7.1).

Major Assumptions Figure 7.1 shows a time line for all three deal scenarios evaluated. Also shown are the milestone schedules for all three scenarios, along with major assumption data. The total time frame for all deal calculations was 307.9 months, where the candidate pharmaceutical gains a 20 percent maximum market share of a 1 billion dollar market, with a 20 percent standard deviation during the projected 15-year sales period of the pharmaceutical. The market is assumed to grow 1.0 percent annually starting at the effective date of the agreement and throughout the valuation period. The manufacturing and marketing costs of the potential pharmaceutical were estimated to be 58 percent, an important assumption considering that royalties are paid on net sales, not gross sales. The total market size, market growth rate, maximum market share, and manufacturing and marketing offset are all Monte Carlo variables following lognormal distributions where

TABLE 7.1 Historical Financial Terms Granted to the Licensor of the Signed Biomedical Collaborative Deal Valued and Optimized in This Case Study

	Deal Scenario			
Component	**Historical**	**Higher-Value, Lower-Risk**	**Higher-Value, Higher-Risk**	**Timing**
Licensing Fees	\$100,000	\$125,000	\$ 85,000	30 days from effective date
Licensing Maintenance Fees	\$100,000	\$125,000	\$ 75,000	First anniversary
	200,000	250,000	150,000	Second anniversary
	300,000	375,000	225,000	Third anniversary
	400,000	500,000	300,000	Fourth anniversary
	500,000	500,000	300,000	Fifth anniversary
R&D Funding	\$250,000	\$275,000	\$165,000	Per year
Milestone Payments	\$500,000	\$660,000	\$910,000	First IND filing in United States or European equivalent
		895,000		Successful conclusion of Phase I clinical trials in the United States or European equivalent
		1,095,000	1,400,000	Successful conclusion of Phase II clinical trials in the United States or European equivalent
	1,500,000	1,375,000	1,650,000	First PLA[a] (or NDA[b]) filing or European equivalent
	4,000,000	1,675,000	1,890,000	NDA approval in the United States or European equivalent
Royalties	2.0% Net Sales	0.5% Net Sales	5.5% Net Sales	

[a]Product license application.

[b]New drug application.

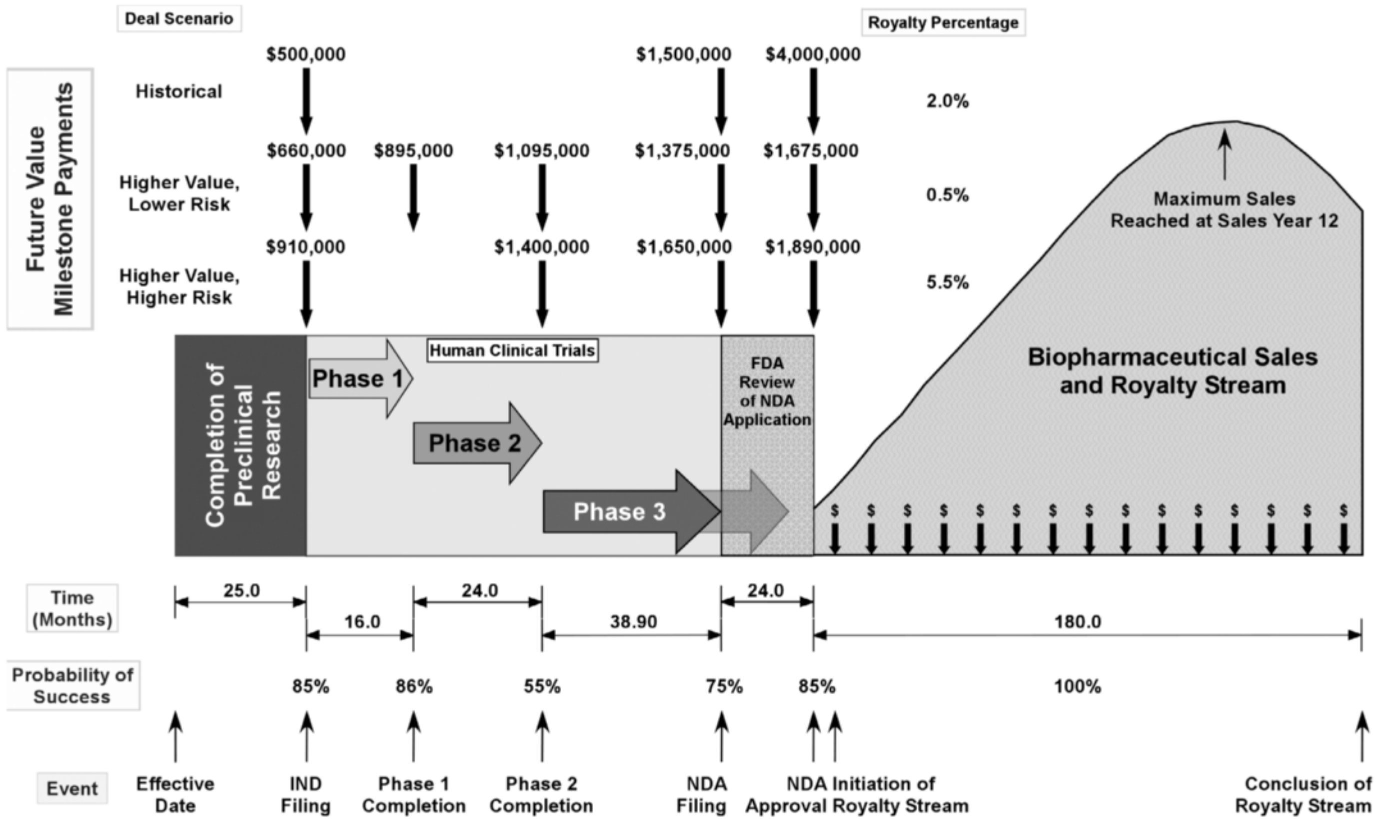


FIGURE 7.1 Time line for the biomedical licensing deal.
Milestone and royalty values for all deal scenarios evaluated are shown. R&D, licensing, and licensing maintenance fees are not shown.

extreme values are unlikely. Assumptions regarding clinical trial length, completion probabilities, and major variables in the valuation model are also shown in Figure 7.1. All of these values are Monte Carlo assumptions. Throughout this case study, deal values were based on royalties from 15 years of net sales. Royalties were paid on a quarterly basis, not at the end of each sales year. Total R&D costs for the licensor were $200,000 annually, again estimated with a Monte Carlo assumption.

Inflation during the period was assumed to be 1.95 percent annually and average annual pharmaceutical price increases (APPIs) were assumed to be 5.8 percent. Thus, milestones were deflated in value, and royalties inflated by APPI less inflation. For the deal valuation described here, the licensor was assumed to be unprofitable preceding and during the clinical trial process and milestone payments were not subject to taxes. However, royalties from the licensee paid to the licensor were taxed at a 33.0 percent rate.

Deal Valuations

Historical Deal Valuation Figure 7.2 illustrates the Monte Carlo summary of the historical deal, while Figure 7.3 shows a comparative illustration of each major component of the historical scenario. Mean deal present value was

Certainty is 50.00% from $1,338,078 to $1,515,976.

Summary
Certainty level is 50.00%.
Certainty range is from $1,338,115 to $1,516,020.
Display range is from $1,091,067 to $1,772,886.
Entire range is from $994,954 to $2,037,413.
After 10,000 trials, the standard error of the mean is $1,344.

Statistics

Trials	10,000
Mean	$1,432,128
Median	$1,422,229
Standard Deviation	$134,449
Variance	$18,076,644,871
Skewness	0.46
Kurtosis	3.47
Coefficient of Variability	9.38%
Range Minimum	$994,954
Range Maximum	$2,037,413
Range Width	$1,042,459
Mean Standard Error	$1,344

FIGURE 7.2 Historical deal scenario Monte Carlo summary.

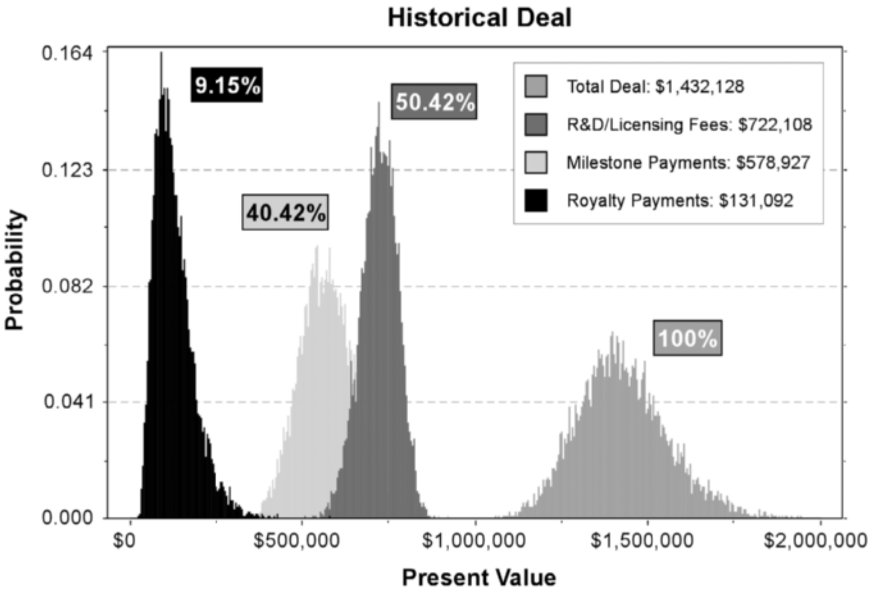


FIGURE 7.3 A comparative illustration I.
This is an illustration of the Monte Carlo distributions of the cash-flow present value of the historical deal scenario, along with the distributions of the deal's individual components. Each component has a clearly definable distribution that differs considerably from other deal components, both in value and risk characteristics. The percentage of each component to total deal present value is also shown.

$1,432,128 with a standard deviation of $134,449 (Figure 7.2). The distribution describing the mean was relatively symmetric with a skewness of 0.46. The kurtosis of the distribution, the "peakedness," was 3.47 (excess kurtosis of 0.47), limiting the deal range from $994,954 to $2,037,413. The coefficient of variation (CV), the primary measure of risk for the deal, was low at 9.38 percent. R&D/licensing contributed the most to total deal value with a mean present value of $722,108, while royalties contributed the least with a mean value of $131,092 (Figure 7.3). Milestones in the historical scenario also contributed greatly to the historical deal value with a mean present value of $578,927.

The riskiness of the cash flows varied greatly among individual historical deal components. R&D/licensing cash flows varied the least and had by far the lowest risk with a CV of only 7.48 percent and, proportional to the distribution's mean, had the smallest range among any deal component (data not shown). The present value of milestone cash flows was much more volatile, with a CV of 14.58 percent. Here the range was greater ($315,103 to $1,004,563) with a symmetric distribution having a skewness of only 0.40 (data not shown).

Royalty present value was by far the most volatile with a CV of 45.71 percent (data not shown). The kurtosis of royalty present value was large (5.98; data not shown), illustrating the proportionally wide distribution to the small royalty mean ($131,093; Figure 7.3). These data should not be surprising as the royalty cash flows are subject to variability of nearly all Monte Carlo assumptions in the model and are thus highly volatile.

Monte Carlo Assumption and Decision Variable Sensitivities Figure 7.4 shows a tornado chart of historical deal assumptions and decision variables. The probability of IND filing had the largest influence on variation of total deal present value, as all milestones and royalties are dependent on this variable. Interestingly, next came the annual research cost for each full-time equivalent (FTE) for the licensor performing the remaining preclinical work in preparation for an IND filing, followed by the negotiated funding amount of each FTE (Figure 7.4). Thus, an area for the licensor to create shareholder value is to overestimate R&D costs in negotiating the financial terms for the deal, considering R&D/licensing funding contributed 50.42 percent of total deal present value (Figure 7.3). Variables impacting royalty cash flows, such as the royalty discount rate and manufacturing and marketing offset percentages, were more important than the negotiated milestone amounts, although the milestone discount rate was 10th in contribution to variance to the historical deal (Figure 7.4).

Higher-Value, Lower-Risk Deal Valuation

Changes in Key Assumptions and Parameters Differing from the Historical, Signed Deal The financial structure for the HVLR deal scenario was considerably different from the historical deal (Table 7.1). Indeed, R&D and licensing funding were significantly increased and the milestone schedule was reorganized with five payments instead of the three in the historical deal. In the HVLR scenario, the value of each individual milestone was stochastically optimized using individual restrictions for each payment. While the future value of the milestone payments was actually $300,000 less than the historical deal (Table 7.1), the present value as determined by Monte Carlo analysis was 93.6 percent higher. In devising this scenario, to compensate the licensee for increased R&D/licensing fees and milestone restructuring, the royalty value in the HVLR scenario was reduced to only a 0.5 percent flat rate (Table 7.1).

Deal Valuation, Statistics, and Sensitivities Figure 7.5 shows the Monte Carlo summary of the HVLR scenario, and Figure 7.6 shows an illustration of present value of the HVLR deal and its three components. The Monte Carlo mean deal value for this scenario was $2,092,617, an increase of 46.1 percent over

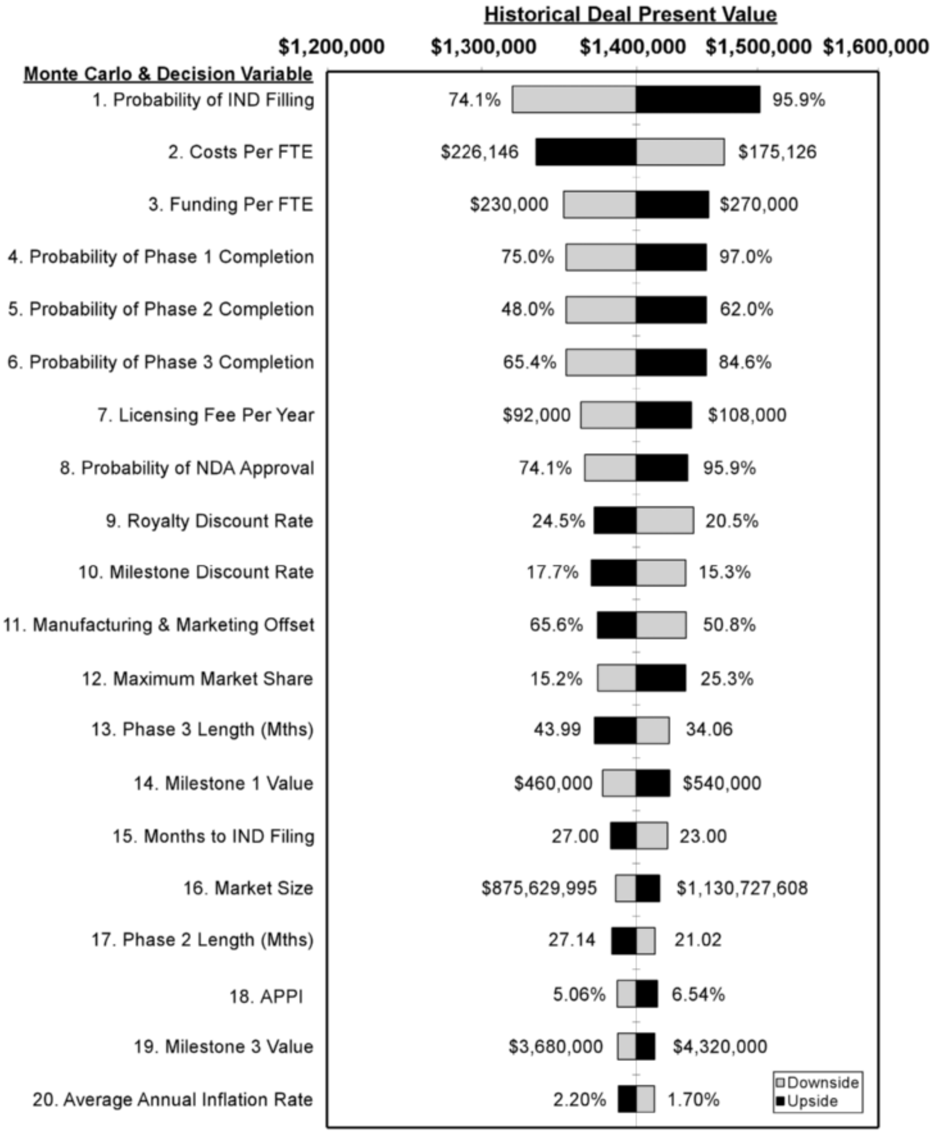


FIGURE 7.4 Historical deal Monte Carlo and decision variable tornado chart.

the historical deal, while total risk was reduced by 16.3 percent as measured by changes in the CV of cash-flow present value (Figures 7.2 and 7.5). This gain in total deal value was achieved by a 93.6 percent increase in the present value of milestone payments (Figures 7.3 and 7.6) along with a 9.6

Certainty is 50.00% from $1,980,294 to $2,200,228.

Summary
Certainty level is 50.00%.
Certainty range is from $1,980,218 to $2,199,958.
Display range is from $1,663,093 to $2,523,897.
Entire range is from $1,475,621 to $2,777,048.
After 10,000 trials, the standard error of the mean is $1,643.

Statistics

Trials	10,000
Mean	$2,092,617
Median	$2,087,697
Standard Deviation	$164,274
Variance	$26,986,218,809
Skewness	0.18
Kurtosis	3.06
Coefficient of Variability	7.85%
Range Minimum	$1,475,620
Range Maximum	$2,777,047
Range Width	$1,301,427
Mean Standard Error	$1,642

FIGURE 7.5 Higher-value, lower-risk deal scenario Monte Carlo.

percent reduction in milestone risk (data not shown). The present value of R&D/licensing funding also increased (30.1 percent) while there is a 22.5 percent reduction in risk. These gains came at the cost of royalty income being reduced by 75.1 percent (Figures 7.3 and 7.6).

The royalty component was so small and the mean so tightly concentrated that the other distributions were comparatively distorted (Panel A, Figure 7.6). If the royalty component is removed, the total deal, milestone, and R&D/licensing distributions are more clearly presented (Panel B, Figure 7.6). The milestone percentage of the total HVLR scenario was much higher than the milestone component of the historical deal, while the R&D/licensing fees of the HVLR structure were less than the historical structure (Figures 7.3 and 7.7).

Cumulatively, the HVLR scenario had a 16.9 percent reduction in risk in comparison to the historical deal (Figures 7.2 and 7.5), where the R&D/licensing and milestone cash flows of HVLR structure were considerably less risky than the historical scenario (data not shown). However, not surprisingly, the risk for the royalty cash flows of the HVLR structure remained nearly identical to that of the historical deal's royalties (data not shown).

Monte Carlo Assumption and Decision Variable Sensitivities The tornado chart for the HVLR deal is presented in Figure 7.8. As with the historical deal, the

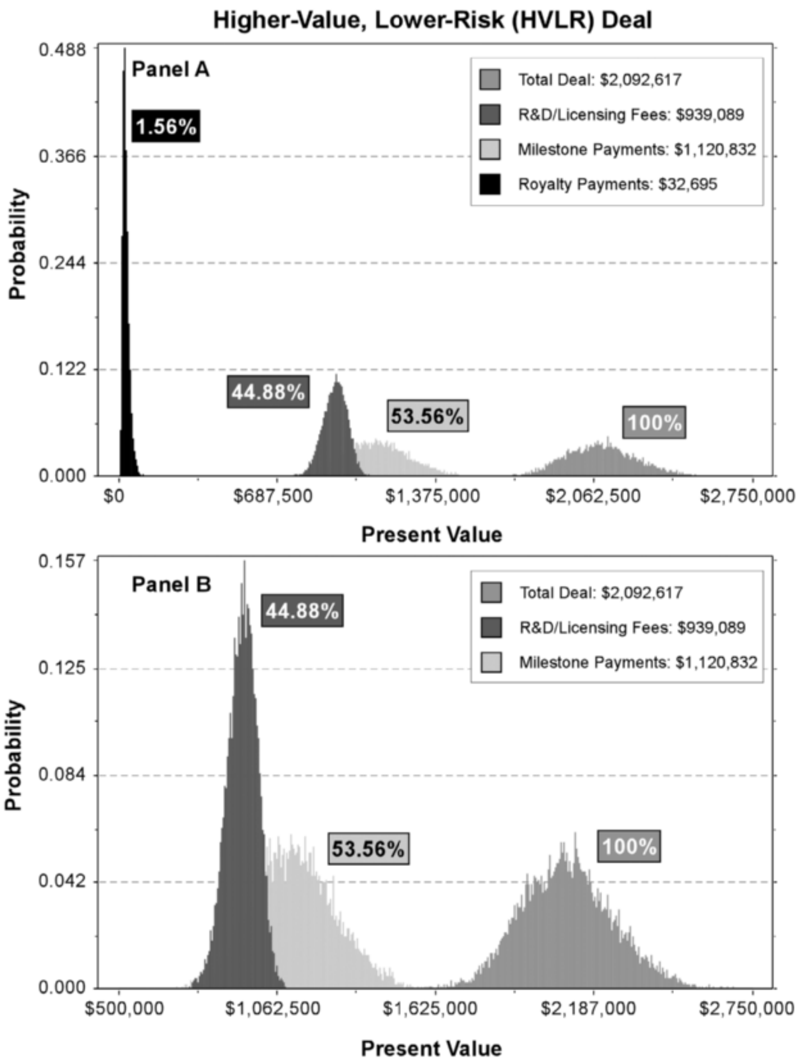


FIGURE 7.6 A comparative illustration II.
The figures illustrate the Monte Carlo distributions for cash-flow present value of the HVLR deal scenario along with the distributions of the deal's individual components. Because the royalty cash flows greatly distort the other distributions (Panel A), removing the royalties from the overlay chart allows the other distributions to be more clearly presented (Panel B). The data in Panel B are comparable to a similar representation of the historical deal (Figure 7.3). Here, proportionally, milestones contributed the most to deal value (53.56 percent), followed by R&D/licensing (44.88 percent), while royalties contributed very little (1.56 percent; Panel A).

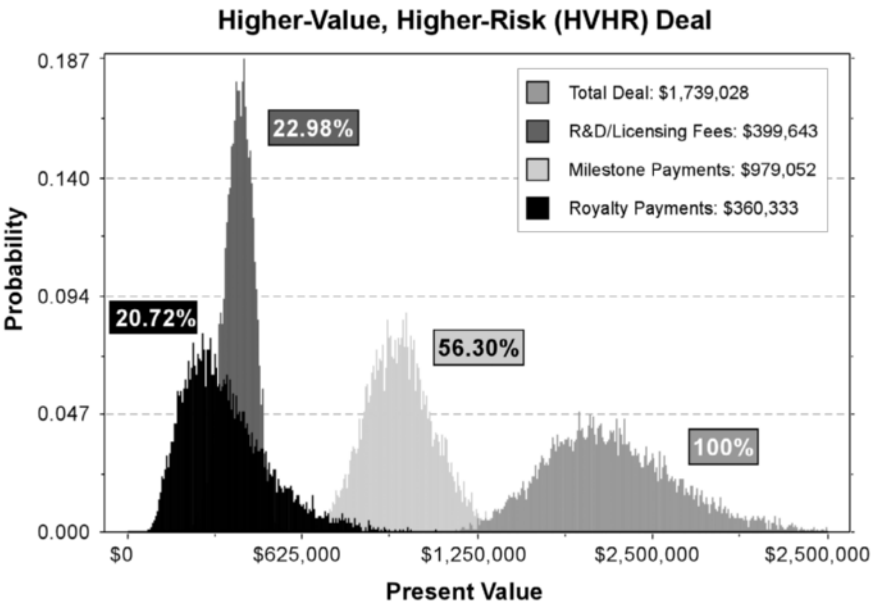


FIGURE 7.7 A comparative illustration III.
Illustrations of the Monte Carlo distributions for cash-flow present value of the HVLR deal scenario along with the distributions of the deal's individual components. Here, proportionally, milestones contributed the most to deal value (56.30 percent), followed by R&D/licensing (22.98 percent), while royalties contributed 20.72 percent to total deal value.

probability of IND filing produced the largest variation in the HVLR deal. The annual research cost for each FTE for the licensor performing the remaining preclinical work in preparation for IND filing was third, while the negotiated annual funding amount for each FTE was fourth. The value of each milestone was listed earlier in importance in comparison to the historical deal (Figures 7.4 and 7.8). This result should not be surprising as the present value of total milestones increased 93.6 percent over the historical structure.

The probabilities of completing various clinical trial stages were not clustered as with the historical deal (Figures 7.4 and 7.8). Indeed, the probability of completing Phase 1 was 2nd, the probability of Phase 2 completion 5th, and the probability of Phase 3 completion 10th in predicting variation in total HVLR deal value (Figure 7.8), whereas in the historical deal, these three variables were clustered and ranked 4th through 6th (Figure 7.4). This reorganization is probably because of milestone restructuring where, in the HVLR deal structure, early milestone payments are worth much more (Table 7.1 and Figure 7.1). Among the top 20 most important

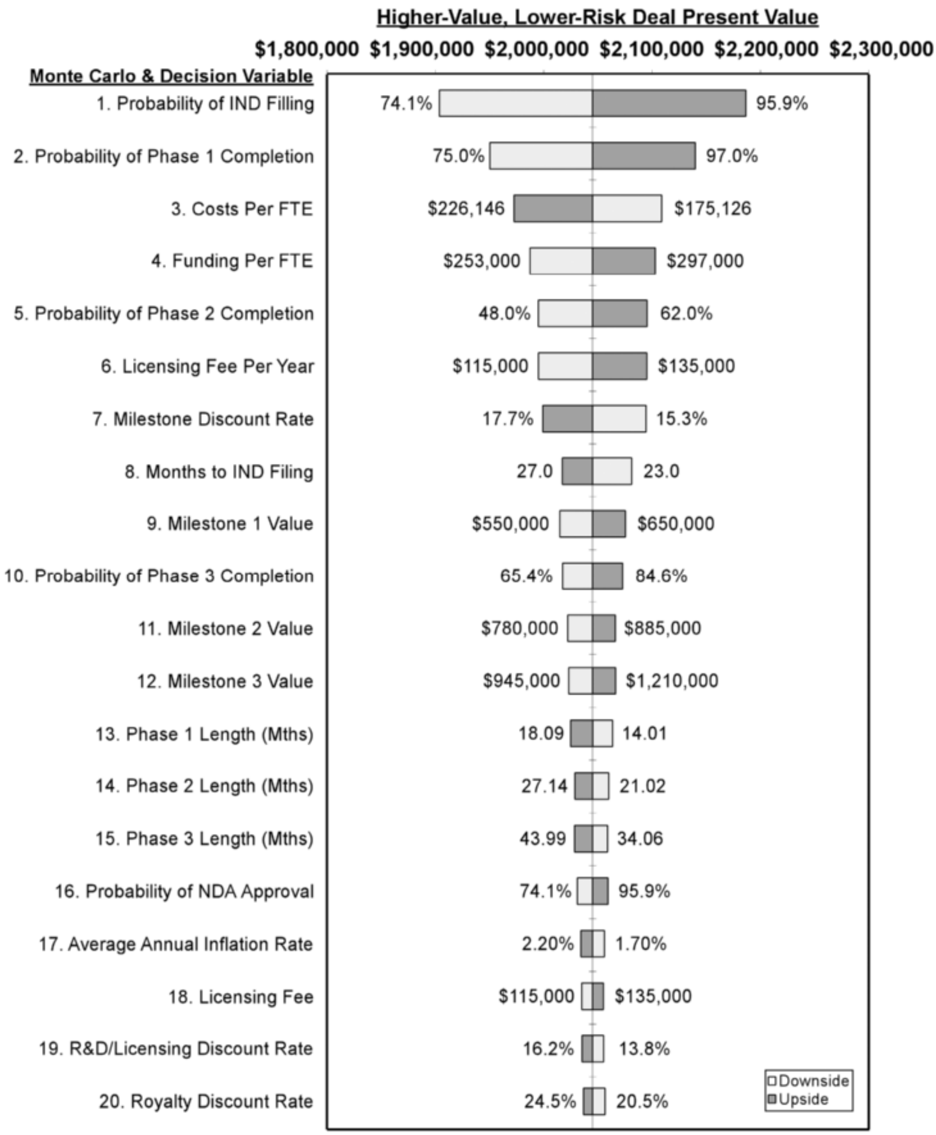


FIGURE 7.8 Higher-value, lower-risk deal scenario Monte Carlo tornado.

variables inducing variation in the HVLR deal are the lengths of Phase 1, Phase 2, and Phase 3 clinical trials (13th–15th; Figure 7.8), although their importance was considerably less than the historical deal (Figure 7.4). This is probably because of the reduced royalty component of the HVLR scenario (Table 7.1).

Higher-Value, Higher-Risk Deal Valuation

Changes in Key Assumptions and Parameters Differing from the Historical and HVLR Deal Structures A variety of financial terms were changed for the HVHR deal structure. First, licensing and licensing maintenance fees were reduced, sometimes substantially (Table 7.1). R&D fees were reduced across the board from the historical deal and the milestone schedule was completely restructured. The historical structure had three payments and the HVLR structure five, with the HVHR deal having only four (Figure 7.1). As shown, the milestone future value for the HVHR deal was reduced to $5,850,000 from $6,000,000 in the historical deal. Like the HVLR deal, the milestone values for the HVHR scenario were stochastically optimized based on specific ranges. The sacrifices gained by lower licensing fees, R&D funding, and milestone restructuring were compensated for by a higher flat royalty rate of 5.5 percent of net sales (Table 7.1).

Deal Valuation, Statistics, and Sensitivities Figure 7.7 shows an illustration of the total HVHR deal along with its three components. Total deal value for the HVHR scenario was $1,739,028, a 21.4 percent increase from the historical deal and 16.9 percent decrease from the HVLR structure. R&D/licensing present value decreased by 44.7 percent and 57.4 percent from the historical and HVLR deals, respectively (Figures 7.3 through 7.7).

The royalty distribution is much more pronounced and noticeably positively skewed, and illustrates the large downside potential of this deal component. Changes in the royalty percentage also significantly expanded the range maximum for the total deal ($3,462,679) with a range width of $2,402,076, a 130.4 percent increase from the historical and 84.6 percent increase over the HVLR deal widths, respectively (Table 7.2).

Milestone present value increased by 69.1 percent from the historical deal and decreased 12.6 percent from the HVLR scenario, while royalty present value increased 175 percent and 1,002 percent, respectively (Figures 7.3 through 7.7). Both the skewness and kurtosis of total deal value under the

TABLE 7.2 Deal Scenario Summary Table as Calculated by Monte Carlo Analysis

Deal Structure	Expected Value	CV	Range Minimum	Range Maximum	Range Width
Historical	$1,432,128	9.38%	$ 994,954	$2,037,413	$1,042,459
Higher-Value, Lower-Risk	2,092,617	7.85	1,475,620	2,777,047	1,301,427
Higher-Value, Higher-Risk	1,739,028	14.33	1,060,603	3,462,679	2,402,076

HVHR scenario were greater than the other deal structures evaluated (Figures 7.3 through 7.7). This result has to do with the greater royalty component in the HVHR scenario and its associated large cash-flow volatility.

The overall deal risk under the HVHR scenario was the greatest (14.33 percent) in comparison to the historical deal's 9.38 percent and the HVLR scenario's 7.85 percent cash-flow CV, again illustrating the strong royalty component of this deal structure with its greater volatility. With the HVHR deal, R&D/licensing cash flows had much higher risk than either the historical or HVLR deals (data not shown). This increased risk is surely because negotiated R&D funding per FTE and licensing fees were considerably less than the estimated cost per FTE, resulting in more R&D/licensing cash-flow volatility in the HVHR structure. This result again shows the importance of accurate accounting and finance in estimating R&D costs for maximizing this type of licensing deal value.

Monte Carlo Assumption and Decision Variable Sensitivities The tornado chart for the HVHR deal scenario emphasized the importance of variables directly impacting royalty cash flows (Figure 7.9). Here, the royalty discount rate was 4th, manufacturing and marketing offset 5th, and maximum market share capture 6th in impacting total deal present value variation. Total market size and the average APPI were 11th and 12th, respectively. Interestingly, the negotiated royalty percentage was only 19th in contribution to deal variance. Cost per FTE ranked 8th, showing this assumption is important in all deal scenarios (Figures 7.4, 7.8, and 7.9). Figure 7.10 shows the Monte Carlo simulation results for HVHR.

The negotiated first milestone value was the only milestone listed on the sensitivity chart (13th, Figure 7.9), illustrating the importance of milestone structuring (Table 7.1 and Figure 7.1). The first milestone is impacted the least by the time value of money and the probability of completion of each clinical trial stage.

A Structural Comparison of Deal Scenario Returns and Risks

Total deal expected value and risk as measured by the CV of cash-flow present value are shown in Table 7.2. As illustrated here, higher expected value is not necessarily correlated with higher risk, which is contrary to a basic principal in finance where investments of higher risk should always yield higher returns. Thus, these data show why quantitative deal valuation and optimization is critical for *all* companies as higher deal values can be constructed with significantly less risk.

Also shown in Table 7.2 are the range minimums, maximums, and widths of the total deal value distributions as calculated by Monte Carlo analysis

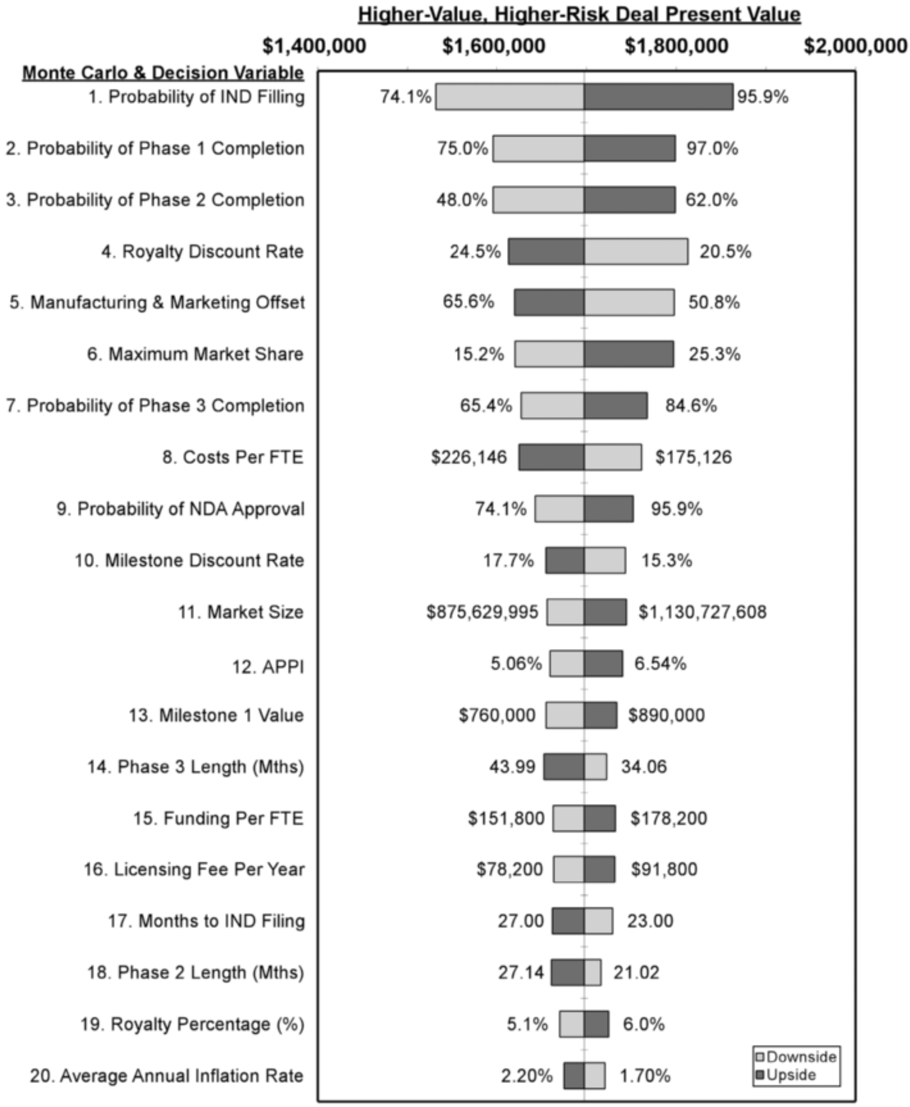


FIGURE 7.9 Higher-value, higher-risk deal scenario Monte Carlo tornado.

for each scenario evaluated. The range minimum is the smallest number and the range maximum the largest number in a distribution, while the range width is the difference between the range minimum and maximum.

Collaborative business deals in the biotechnology and pharmaceutical industries formed during strategic alliances, such as the one described here, are

Certainty is 50.00% from $1,563,891 to $1,882,975.

Summary
Certainty level is 50.00%.
Certainty range is from $1,563,891 to $1,882,975.
Display range is from $1,132,837 to $2,396,924.
Entire range is from $1,060,603 to $3,462,679.
After 10,000 trials, the standard error of the mean is $2,493.

Statistics

Trials	10,000
Mean	$1,739,028
Median	$1,712,532
Standard Deviation	$249,257
Variance	$62,129,317,618
Skewness	0.77
Kurtosis	4.39
Coefficient of Variability	14.33%
Range Minimum	$1,060,603
Range Maximum	$3,462,679
Range Width	$2,402,076
Mean Standard Error	$2,492

FIGURE 7.10 Higher-value, higher-risk deal scenario Monte Carlo summary.

in fact risky asset portfolios. As such, the standard deviation of a portfolio of assets is less than the weighted average of the component asset standard deviations. To view the impact of diversification of cash-flow streams with the various deal scenarios evaluated in this case study, the weight of each deal component was determined and the weighted average CV of cash-flow present value calculated for each deal scenario (Table 7.3). The CV is used as the primary risk measure because of differences in the scale of the cash flows from individual deal components.

As expected with a portfolio of risky assets, the weighted average of the CV of individual deal components (R&D/licensing funding, milestone payments, and royalties) was always greater than the CV of the total deal present value, illustrating the impact of diversification (Table 7.3). Thus, portfolios of less than perfectly correlated assets always offer better risk–return opportunities than the individual component assets on their own. As such, companies would probably not want to completely forgo receiving milestone payments and royalties for only R&D funding and licensing fees, *if* these deal components can be valued and optimized with reasonable accuracy as described here. By combining assets whose returns are uncorrelated or partially correlated, such as cash flows from milestone payments, royalties, licensing, and R&D funding, risk is reduced (Table 7.3). Risk can be eliminated most rapidly while keeping expected returns as high as possible if a

TABLE 7.3 Deal Component Weights, Component CVs, Weighted Average Deal CVs, and Calculated Deal CVs

	Weights			Coefficient of Variation (CV)				
Deal Structure	$W_{R\&D}$[a]	W_{Mi}[b]	W_{Ry}[c]	**R&D**[d]	**Milestones**	**Royalties**	**W. Avg.**[e]	**Calculated**[f]
Historical	50.42%	40.42%	9.17%	7.47%	14.57%	45.70%	13.84%	9.38%
Higher-Value, Lower-Risk	44.88	53.56	1.56	5.79	13.18	45.95	10.38	7.85
Higher-Value, Higher-Risk	22.98	56.30	20.72	13.40	12.69	46.21	19.80	14.33

[a]Proportion of total deal present value attributable to R&D and licensing fees.

[b]Proportion of total deal present value attributable to milestone payments.

[c]Proportion of total deal present value attributable to royalty payments.

[d]CV in the present value of cash flows from R&D and licensing fees.

[e]Weighted average of the CV of total deal value.

[f]Calculated deal CV by Monte Carlo simulation.

company's cumulative deal repertoire is valued, structured, and balanced from the beginning of a company's evolution and development.

Discussion and Conclusion

The historical deal evaluated in this case study was a preclinical, product-licensing deal for a biopharmaceutical with one major therapeutic indication. For collaborative deal structures containing licensing fees, R&D funding, milestone payments, and royalties, each deal component has definable expected values, variances, and widely varying risk characteristics. Alternative deal structures were developed and optimized, all of which had different expected returns and risk levels with the primary risk measure being the CV of cash-flow present values. Thus, nearly any biomedical collaborative deal with the types of financial terms described here can be quantitatively valued, structured, and optimized using financial models, Monte Carlo analysis, stochastic optimization, real options, and portfolio theory.

During this study, the author was at a considerable disadvantage because the historical deal valued and optimized here had already been signed, and he was not present during the negotiation process. Therefore, the author had to make a large number of assumptions when restructuring the financial terms of the agreement. Considering these limitations, this case is not about what is appropriate in the comparative financial terms for a biomedical licensing deal and what is not; rather, the data described here are valuable in showing the quantitative influence of different deal structures on the overall valuation of a biomedical collaborative agreement, and most importantly on the level of overall deal risk, as well as the risk of the individual deal components. The most effective approach using this technique is to work with a negotiator during the development and due diligence, and through the closing process of a collaborative agreement. During this time, data should be continually gathered and the financial models refined as negotiations and due diligence proceed.

CASE STUDY: OIL AND GAS EXPLORATION AND PRODUCTION

This case study was contributed by Steve Hoye. Steve is an independent business consultant with more than 23 years of oil and gas industry experience, specializing in Monte Carlo simulation for the oil and gas industry. Starting with a bachelor of science degree from Purdue University in 1980, he served as a geophysicist with Texaco in Houston, Denver, and Midland, Texas, before earning the MBA degree from the University of Denver in 1997. Since then, Steve has held leadership roles with Texaco as the midcontinent

BU technology team leader, and as asset team manager in Texaco's Permian Basin business unit, before starting his consultancy in 2002. Steve can be reached at steve@hoyeconsultinggroup.com.

The oil and gas industry is an excellent place to examine and discuss techniques for analyzing risk. The basic business model discussed involves making investments in land rights, geologic data, drilling (services and hardware), and human expertise in return for a stream of oil or gas production that can be sold at a profit. This model is beset with multiple, significant risk factors that determine the resulting project's profitability, including:

- *Dry-Hole Risk*. Investing drilling dollars with no resulting revenue from oil or gas because none is found in the penetrated geologic formation.
- *Drilling Risk*. High drilling costs can often ruin a project's profitability. Although companies do their best to estimate them accurately, unforeseeable geological or mechanical difficulties can cause significant variability in actual costs.
- *Production Risk*. Even when oil or gas reservoirs are discovered by drilling, there is a high probability that point estimates of the size and recoverability of the hydrocarbon reserves over time are wrong.
- *Price Risk*. Along with the cyclical nature of the oil and gas industry, product prices can also vary unexpectedly during significant political events such as war in the Middle East, overproduction and cheating by the OPEC cartel, interruptions in supply such as large refinery fires, labor strikes, or political uprisings in large producing nations (e.g., Venezuela in 2002), and changes in world demand.
- *Political Risk*. Significant amounts of the world's hydrocarbon reserves are controlled by nations with unstable governments. Companies that invest in projects in these countries take significant risks that the governments and leaders with whom they have signed contracts will no longer be in power when earned revenue streams should be shared contractually. In many well-documented cases, corporate investments in property, plants, and equipment (PPE) are simply nationalized by local governments, leaving companies without revenue or the equipment and facilities that they built to earn that revenue.

Oil and gas investments generally are very capital-intensive, often making these risks more than just of passing interest. Business units and entire companies stake their survival on their ability to properly account for these risks as they apportion their capital budgets in a manner that ensures value to their stakeholders. To underline the importance of risk management in the industry, many large oil companies commission high-level corporate panels of experts to review and endorse risk assessments done across all of their business units for large capital projects. These reviews attempt to ensure consistency of risk assessment across departments and divisions that are often under pressure to make their investment portfolios look attractive to corporate leadership as they compete for capital.

Monte Carlo simulation is a preferred approach to the evaluation of the multiple, complex risk factors in the model we discuss. Because of the inherent complexity of these risk factors and their interactions, deterministic solutions are not practical, and point forecasts are of limited use and, at worst, are misleading. In contrast, Monte Carlo simulation is ideal for economic evaluations under these circumstances. Domain experts can individually quantify and describe the project risks associated with their areas of expertise without having to define their overall effect on project economics.[1] Cash-flow models that integrate the diverse risk assumptions for each of the prospect team's experts are relatively straightforward to construct and analyze. Most importantly, the resulting predictions of performance do not result in a simple single-point estimate of the profitability of a given oil and gas prospect. Instead, they provide management with a spectrum of possible outcomes and their related probabilities. Best of all, Monte Carlo simulation provides estimates of the sensitivities of their investment outcomes to the critical assumptions in their models, allowing them to focus money and people on the critical factors that will determine whether they meet the financial goals defined in their business plans. Ultimately, Monte Carlo simulation becomes a project management tool that decreases risk while increasing profits.

In this case study, we explore a practical model of an oil-drilling prospect, taking into account many of the risk factors described earlier. While the model is hypothetical, the general parameters we use are consistent with those encountered drilling in a mature, oil-rich basin in the United States (e.g., Permian Basin of West Texas) in terms of the risk factors and related revenues and expenses. This model is of greater interest as a framework and approach than it is as an evaluation of any particular drilling prospect. Its value is in demonstrating the approach to quantifying important risk assumptions in an oil prospect using Monte Carlo simulation, and analyzing their effects on the profitability forecasts of the project. The techniques described herein are extensible to many other styles and types of oil and gas prospects.

Cash-Flow Model

The model was constructed using Risk Simulator, which provides all of the necessary Monte Carlo simulation tools as an easy-to-use, comprehensive add-in to Microsoft Excel. The model simulates the drilling outcome as being a dry-hole or an oil discovery using dry-hole risk factors for the particular geologic formation and basin. Drilling, seismic, and land-lease costs are incurred whether the well is dry or a discovery. If the well is a discovery, a revenue stream is computed for the produced oil over time using assumptions for product price, and for the oil production rate as it declines over time from its initial value. Expenses are deducted for royalty payments to landowners, operating costs associated with producing the oil, and severance taxes levied by states on the produced oil. Finally, the resulting net cash flows are discounted at the weighted average cost of capital (WACC) for the firm and summed to a net present value (NPV) for the project. Each of these sections of the model is now discussed in more detail.

Dry-Hole Risk

Companies often have proprietary schemes for quantifying the risk associated with not finding any oil or gas in their drilled well. In general, though, there are four primary and independent conditions that must all be encountered in order for hydrocarbons to be found by the drill bit:

1. *Hydrocarbons* must be present.
2. A *reservoir* must be developed in the rock formation to hold the hydrocarbons.
3. An impermeable *seal* must be available to trap the hydrocarbons in the reservoir and prevent them from migrating somewhere else.
4. A *structure* or *closure* must be present that will cause the hydrocarbons (sealed in the reservoir) to pool in a field where the drill bit will penetrate.

Because these four factors are independent and must each be true in order for hydrocarbons to be encountered by the drill bit (and a dry hole to be avoided), the probability of a producing well is defined as:

$$P_{\text{Producing Well}} = P_{\text{Hydrocarbons}} \times P_{\text{Reservoir}} \times P_{\text{Seal}} \times P_{\text{Structure}}$$

Figure 7.11 shows the model section labeled “Dry-Hole Risk,” along with the probability distributions for each factor’s Monte Carlo assumption. While a project team most often describes each of these factors as a single-point estimate, other methods are sometimes used to quantify these risks. The most effective process the author has witnessed involved the presentation of the geological, geophysical, and engineering factors by the prospect team to a group of expert peers with wide experience in the proposed area. Thes peer experts then rated each of the risk factors. The resulting distribution of risk factors often appeared near-normally distributed, with strong central tendencies and symmetrical tails. This approach was very amenable

Dry-Hole Risk Risk Factor	Prob. of Success	Mean	Stdev	Min	Max
Hydrocarbons	89.7%	99.0%	5.0%	0	100%
Structure	89.7%	100.0%	0.0%	0	100%
Reservoir	89.7%	75.0%	10.0%	0	100%
Seal	89.7%	100.0%	0.0%	0	100%
Net Producing Well Prob.:	**64.8%**				
Producing Well [0=no,1=yes]	1				

FIGURE 7.11 Dry-hole risk.

to Monte Carlo simulation. It highlighted those factors where there was general agreement about risk and brought the riskiest factors to the foreground where they were examined and specifically addressed.

Accordingly, the assumptions regarding dry-hole risk in this model reflect a relatively low risk profile.[2] Each of the four risk factor assumptions in Figure 7.11 (dark shaded area) are described as normally distributed variables, with the mean and standard deviations for each distribution to the right of the assumption fields. The ranges of these normal distributions are confined and truncated between the *min* and *max* fields, and random samples for any simulation trial outside this range are ignored as unrealistic.

As described earlier, the *Net Producing Well Probability* field in the model corresponds to the product of the four previously described risk factors. These four risk factors are drawn as random samples from their respective normal distributions for each trial or iteration of the simulation. Finally, as each iteration of the Monte Carlo simulation is conducted, the field labeled *Producing Well* generates a random number between zero and one to determine if that simulation resulted in a discovery of oil or a dry hole. If the random number is less than the *Net Producing Well Probability*, it is a producing well and shows the number one. Conversely, if the random number is greater than the *Net Producing Well Probability*, the simulated well is a dry hole and shows zero.

Production Risk

A multiyear stream of oil can be characterized as an initial oil production rate (measured in barrels of oil per day, BOPD), followed by a decline in production rates as the natural reservoir energy and volumes are depleted over time. Reservoir engineers can characterize production declines using a wide array of mathematical models, choosing those that most closely match the geology and producing characteristics of the reservoir. Our hypothetical production stream is described with two parameters:

1. *IP*. The initial production rate tested from the drilled well.
2. *Decline Rate*. An exponentially declining production rate that describes the annual decrease in production from the beginning of the year to the end of the same year. Production rates in BOPD for our model are calculated by:

$$\text{Rate}_{\text{Year End}} = (1 - \text{Decline Rate}) \times \text{Rate}_{\text{Year Begin}}$$

Yearly production volumes in barrels of oil are approximated as:

$$\text{Oil Volume}_{\text{Year}} = 365 \times (\text{Rate}_{\text{Year Begin}} + \text{Rate}_{\text{Year End}})/2$$

For Monte Carlo simulation, our model represents the IPs with a lognormal distribution with a mean of 441 BOPD and a standard deviation of 165 BOPD. The decline rate was modeled with a uniform probability of occurrence between 15 percent and 28 percent. To add interest and realism to our hypothetical model, we incorporated an additional constraint in the production model that simulates a situation that might occur for a particular reservoir where higher IPs imply that the production decline rate will be higher. This constraint is implemented by imposing a correlation coefficient of 0.60 between the IP and decline rate assumptions that are drawn from their respective distributions during each trial of the simulation.

The production and operating expense sections of the model are shown in Figure 7.12. Although only the first 3 years are shown, the model accounts for up to 25 years of production. However, when production declines below the economic limit,[3] it will be zeroed for that year and every subsequent year, ending the producing life of the well. As shown, the IP is assumed

		End of Year:			
	Decline Rate	0	1	2	3
BOPD	21.5%	442	347	272	214
Net BBLS / Yr			143,866	112,924	88,636
Price / BBl			$ 20.14	$ 20.14	$ 20.14
Net Revenue Interest	77.4%		77.4%	77.4%	77.4%
Revenue			**$ 2,242,311**	**$ 1,760,035**	**$ 1,381,487**
Operating Costs [$/Barrel]	$ 4.80		$ (690,558)	$ (542,033)	$ (425,453)
Severance Taxes [$]	6.0%	rate	$ (134,539)	$ (105,602)	$ (82,889)
Net Sales			$ 1,417,214	$ 1,112,400	$ 873,145

FIGURE 7.12 Decline rate.

to occur at the end of Year 0, with the first full year of production accounted for at the end of Year 1.

Revenue Section

Revenues from the model flow literally from the sale of the oil production computed earlier. Again there are two assumptions in our model that represent risks in our prospect:

1. *Price.* Over the past 10 years, oil prices have varied from $13.63/barrel in 1998 to nearly $30/barrel in 2000.[4] Consistent with the data, our model assumes a normal price distribution with a mean of $20.14 and a standard deviation of $4.43/barrel.
2. *Net Revenue Interest.* Oil companies must purchase leases from mineral interest holders. Along with paying cash to retain the drilling and production rights to a property for a specified time period, the lessee also generally retains some percentage of the oil revenue produced in the form of a royalty. The percentage that the producing company retains after paying all royalties is the net revenue interest (NRI). Our model represents a typical West Texas scenario with an assumed NRI distributed normally with a mean of 75 percent and a standard deviation of 2 percent.

The revenue portion of the model is also shown in Figure 7.12 immediately below the production stream.

The yearly production volumes are multiplied by sampled price per barrel, and then multiplied by the assumed NRI to reflect dilution of revenues from royalty payments to lessees.

Operating Expense Section

Below the revenue portion are operating expenses, which include two assumptions:

1. *Operating Costs.* Companies must pay for manpower and hardware involved in the production process. These expenses are generally described as a dollar amount per barrel. A reasonable West Texas cost would be $4.80 per barrel with a standard deviation of $0.60 per barrel.
2. *Severance Taxes.* State taxes levied on produced oil and gas are assumed to be a constant value of 6 percent of revenue.

Operating expenses are subtracted from the gross sales to arrive at net sales, as shown in Figure 7.12.

Drilling Costs	$ 1,209,632
Completion Cost	$ 287,000
Professional Overhead	$ 160,000
Lease Costs / Well	$ 469,408
Seismic Costs / Well	$ 81,195

FIGURE 7.13 Year 0 expenses.

Year 0 Expenses Figure 7.13 shows the Year 0 expenses assumed to be incurred before oil production from the well (and revenue) is realized. These expenses are:

1. *Drilling Costs*. These costs can vary significantly as previously discussed, due to geologic, engineering, and mechanical uncertainty. It is reasonable to skew the distribution of drilling costs to account for a high-end tail consisting of a small number of wells with very large drilling costs due to mechanical failure and unforeseen geologic or serendipitous occurrences. Accordingly, our distribution is assumed to be lognormal, with a mean of $1.2 million and a standard deviation of $200,000.
2. *Completion Costs*. If it is determined that there is oil present in the reservoir (and we have not drilled a dry hole), engineers must prepare the well (mechanically/chemically) to produce oil at the optimum sustainable rates.[5] For this particular well, we hypothesize our engineers believe this cost is normally distributed with a mean of $287,000 and a standard deviation of $30,000.
3. *Professional Overhead*. This project team costs about $320,000 per year in salary and benefits, and we believe the time they have spent is best represented by a triangular distribution, with a most likely percentage of time spent as 50 percent, with a minimum of 40 percent, a maximum of 65 percent.
4. *Seismic and Lease Costs*. To develop the proposal, our team needed to purchase seismic data to choose the optimum well location, and to purchase the right to drill on much of the land in the vicinity of the well. Because this well is not the only well to be drilled on this seismic data and land, the cost of these items is distributed over the planned number of wells in the project. Uncertain assumptions are shown in Figure 7.14, and include leased acres, which were assumed to be normally distributed with a mean of 12,000 and a standard deviation of 1,000 acres. The total number of planned wells over which to distribute the costs was assumed to be uniform between 10 and 30. The number of seismic sections acquired was also assumed to be normally distributed with a mean

Lease Expense		**Comments**
Project Lease Acres	12,800	20 sections
Planned Wells	20.0	
Acres / Well	640	
Acreage Price	$ 733.45	$ / acre
Acreage Cost / Well	**$ 469,408**	

Seismic Expense		
Seismic Sections Acquired	50.0	
Seismic Sections / Well	2.50	
Seismic Cost	$ 32,478.18	$ / section
Seismic Cost / Well	**$ 81,195**	

FIGURE 7.14 Uncertain assumptions.

of 50 sections and a standard deviation of 7. These costs are represented as the final two lines of Year 0 expenses in Figure 7.13.

Net Present Value Section

The final section of the model sums all revenues and expenses for each year starting at Year 0, discounted at the weighted average cost of capital (WACC—which we assume for this model is 9 percent per year), and summed across years to compute the forecast of NPV for the project. In addition, NPV/I is computed,[6] as it can be used as a threshold and ranking mechanism for portfolio decisions as the company determines how this project fits with its other investment opportunities given a limited capital budget.

Monte Carlo Simulation Results

As we assess the results of running the simulation with the assumptions defined previously, it is useful to define and contrast the point estimate of project value computed from our model using the mean or most likely values of the earlier assumptions. The expected value of the project is defined as:

$$\begin{aligned} E_{\text{Project}} &= E_{\text{Dry Hole}} + E_{\text{Producing Well}} \\ &= P_{\text{Dry Hole}} NPV_{\text{Dry Hole}} + P_{\text{Producing Well}} NPV_{\text{Producing Well}} \end{aligned}$$

where $P_{\text{Producing Well}}$ = probability of a producing well and $P_{\text{Dry Hole}}$ = probability of a dry hole = $(1 - P_{\text{Producing Well}})$. Using the mean or most likely point estimate values from our model, the expected NPV of the project is $1,250,000, which might be a very attractive prospect in the firm's portfolio.

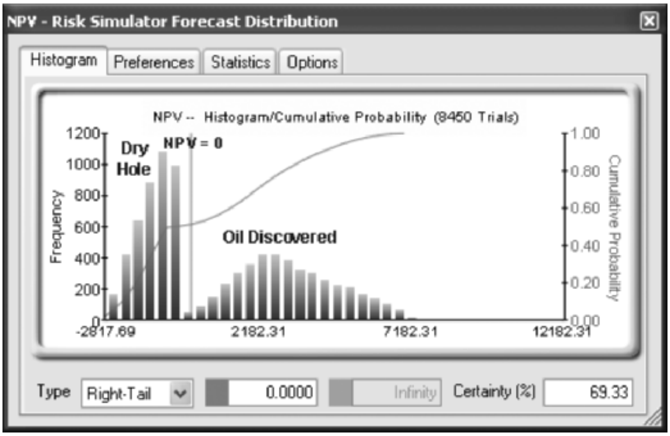


FIGURE 7.15 Frequency distribution of NPV outcomes.

In contrast, we can now examine the spectrum of outcomes and their probability of occurrence. Our simulation was run with 8,450 trials (trial size selected by precision control) to forecast NPV, which provided a mean NPV plus or minus $50,000 with 95 percent confidence. Figure 7.15 is the frequency distribution of NPV outcomes. The distribution is obviously bimodal, with the large, sharp negative NPV peak to the left representing the outcome of a dry hole. The smaller, broader peak toward the higher NPV ranges represents the wider range of more positive NPVs associated with a producing well.

All negative NPV outcomes are to the left of the *NPV = 0* line (with a lighter shade) in Figure 7.15, while positive outcome NPVs are represented by the area to the right of the *NPV = 0* line with the probability of a positive outcome (breakeven or better) shown as 69.33 percent. Of interest, the negative outcome possibilities include not only the dry-hole population of outcomes as shown, but also a small but significant portion of producing-well outcomes that could still lose money for the firm. From this information, we can conclude that there is a 30.67 percent chance that this project will have a negative NPV.

It is obviously not good enough for a project of this sort to avoid a negative NPV. The project must return to shareholders something higher than its cost of capital, and, further, must be competitive with other investment opportunities that the firm has. If our hypothetical firm had a hurdle rate of NPV/I greater than 25 percent for its yearly budget, we would want to test our simulated project outcomes against the probability that the project could clear that hurdle rate.

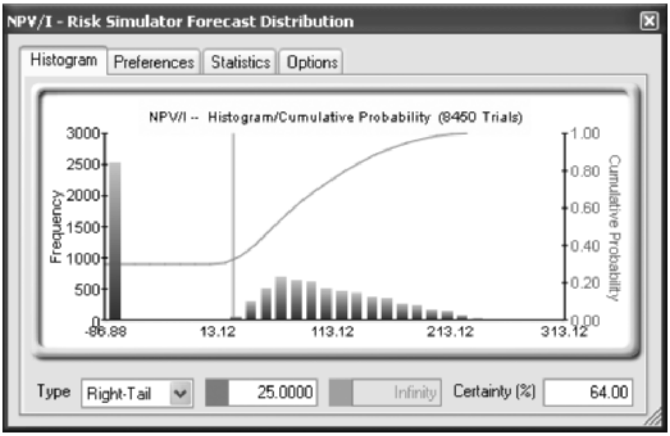


FIGURE 7.16 Forecast distribution of NPV to I ratio.

Figure 7.16 shows the forecast distribution of outcomes for NPV/I. The large peak at negative 100 percent again represents the dry-hole case, where in fact the NPV of the outcome is negative in the amount of Year 0 costs incurred, making NPV/I equal to –1. All outcomes for NPV greater than the hurdle rate of 25 percent show that there is a 64 percent probability that the project will exceed that rate. To a risk-sensitive organization, this outcome implies a probability of greater than one in three that the project will fail to clear the firm's hurdle rate—significant risk indeed.

Finally, our simulation gives us the power to explore the sensitivity of our project outcomes to the risks and assumptions that have been made by our experts in building the model. Figure 7.17 shows a sensitivity analysis of the NPV of our project to the assumptions made in our model. This chart shows the correlation coefficient of the top 10 model assumptions to the NPV forecast in order of decreasing correlation.

At this point, the project manager is empowered to focus resources on the issues that will have an impact on the profitability of this project. Given the information from Figure 7.17, we could hypothesize the following actions to address the top risks in this project in order of importance:

- *IP*. The initial production rate of the well has a driving influence on value of this project, and our uncertainty in predicting this rate is causing the largest swing in predicted project outcomes. Accordingly, we could have our team of reservoir and production engineers further examine known production IPs from analogous reservoirs in this area, and perhaps attempt to stratify the data to further refine predictions of

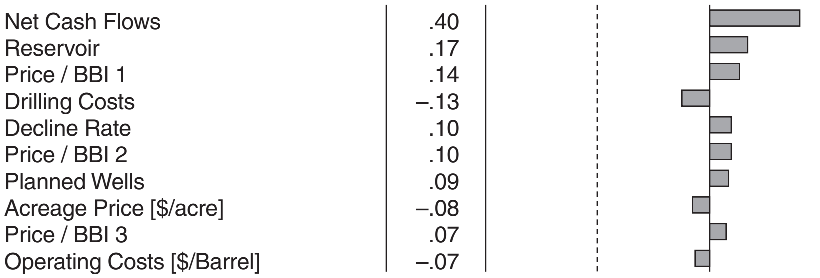


FIGURE 7.17 NPV sensitivity analysis.

IPs based on drilling or completion techniques, geological factors, or geophysical data.

- *Reservoir Risk*. This assumption is the driver of whether the well is a dry hole or producer, and as such it is not surprising that it is a major driving factor. Among many approaches, the project team could investigate the possibility that inadequate analysis of subsurface data is causing many companies to declare dry holes in reservoirs that have hidden producing potential.
- *Oil Price (Year 1) and Drilling Costs*. Both of these items are closely related in their power to affect NPV. Price uncertainty could best be addressed by having a standard price prediction for the firm against which all projects would be compared.[7] Drilling costs could be minimized by process improvements in the drilling team that would tighten the variation of predicted costs from actual costs. The firm could seek out companies with strong track records in their project area for reliable, low-cost drilling.
- *Decline Rate*. The observant reader will note a positive-signed correlation between decline rate and project NPV. At first glance this is unexpected, because we would normally expect that a higher decline rate would reduce the volumes of oil to be sold and hurt the revenue realized by our project. Recall, however, that we correlated higher IPs with higher decline rates in our model assumptions, which is an indirect indication of the power of the IP on the NPV of our project: Despite higher decline rates, the positive impact of higher IPs on our project value is overriding the lost production that occurs because of the rapid reservoir decline. We should redouble our efforts to better predict IPs in our model.

Conclusion

Monte Carlo simulation can be an ideal tool for evaluating oil and gas prospects under conditions of significant and complex uncertainty in the assumptions that would render any single-point estimate of the project outcome nearly useless. The technique provides each member of multidisciplinary work teams a straightforward and effective framework for quantifying and accounting for each of the risk factors that will influence the outcome of his or her drilling project. In addition, Monte Carlo simulation provides management and team leadership something much more valuable than a single forecast of the project's NPV: It provides a probability distribution of the entire spectrum of project outcomes, allowing decision makers to explore any pertinent scenarios associated with the project value. These scenarios could include break-even probabilities as well as scenarios associated with extremely poor project results that could damage the project team's credibility and future access to capital, or outcomes that resulted in highly successful outcomes. Finally, Monte Carlo simulation of oil and gas prospects provides managers and team leaders critical information on which risk factors and assumptions are driving the projected probability of project outcomes, giving them the all-important feedback they need to focus their people and financial resources on addressing those risk assumptions that will have the greatest positive impact on their business, improving their efficiency and adding profits to their bottom line.

CASE STUDY: FINANCIAL PLANNING WITH SIMULATION

Tony Jurado is a financial planner in northern California. He has a BA from Dartmouth College and is a candidate for the Certified Financial Planner designation. Tony specializes in the design and implementation of comprehensive financial plans for high-net-worth individuals. He can be contacted at tony.jurado@alum.dartmouth.org.

Corporate America has increasingly altered the retirement landscape by shifting from defined benefit to defined contribution plans. As the baby boomers retire, they will have different financial planning needs than those of previous generations because they must manage their own retirement funds. A thoughtful financial planner has the ability to positively impact the lives of these retirees.

A Deterministic Plan

Today was the last day of work for Henry Tirement, and, until just now, he and his financial planner, Mr. Determinist, had never seriously discussed what to do with his 401k rollover. After a moment of fact gathering with Henry, Mr. D obtains the following information:

- Current assets are $1,000,000 in various mutual funds.
- Current age is 65.
- Desired retirement salary is $60,000 before-tax.
- Expected return on investments is 10 percent.
- Expected inflation is 3 percent.
- Life expectancy is age 95.
- No inheritance considerations.

With his financial calculator, Mr. D concludes that Henry can meet his retirement goals and, in fact, if he died at age 95, he'd have over $3.2 million in his portfolio. Mr. D knows that past performance does not guarantee future results, but past performance is all that we have to go by. With the stock market averaging over 10 percent for the past 75 years, Mr. D feels certain that this return is reasonable. As inflation has averaged 3 percent over the same time period, he feels that this assumption is also realistic. Mr. D delivers the good news to Henry and the plan is put into motion (Table 7.4).

Fast forward to 10 years later. Henry is not so thrilled anymore. He visits the office of Mr. D with his statements in hand and they sit down to discuss the portfolio performance. Writing down the return of each of the past 10 years, Mr. D calculates the average performance of Henry's portfolio (Table 7.5).

"You've averaged 10 percent per year!" Mr. D tells Henry. Befuddled, Henry scratches his head. He shows his last statement to Mr. D that shows a portfolio balance is $501,490.82.

Once again, Mr. D uses his spreadsheet program and obtains the results in Table 7.6.

Mr. D is not certain what has happened. Henry took out $60,000 at the beginning of each year and increased this amount by 3 percent annually. The portfolio return averaged 10 percent. Henry should have over $1.4 million by now.

Sequence of Returns Sitting in his office later that night, Mr. D thinks hard about what went wrong in the planning. He wonders what would have happened if the annual returns had occurred in reverse order (Table 7.7). The average return is still 10 percent and the withdrawal rate has not changed, but the portfolio ending balance is now $1.4 million. The only difference between the two situations is the sequence of returns. Enlightenment overcomes Mr. D, and he realizes that he has been employing a deterministic planning paradigm during a period of withdrawals.

TABLE 7.4 The Deterministic Plan

Year	Returns (%)	Beginning Balance ($)	Withdrawal ($)	Ending Balance ($)
1	10.00	1,000,000.00	60,000.00	1,034,000.00
2	10.00	1,034,000.00	61,800.00	1,069,420.00
3	10.00	1,069,420.00	63,654.00	1,106,342.60
4	10.00	1,106,342.60	65,563.62	1,144,856.88
5	10.00	1,144,856.88	67,530.53	1,185,058.98
6	10.00	1,185,058.98	69,556.44	1,227,052.79
7	10.00	1,227,052.79	71,643.14	1,270,950.62
8	10.00	1,270,950.62	73,792.43	1,316,874.01
9	10.00	1,316,874.01	76,006.20	1,364,954.58
10	10.00	1,364,954.58	78,286.39	1,415,335.01
11	10.00	1,415,335.01	80,634.98	1,468,170.03
12	10.00	1,468,170.03	83,054.03	1,523,627.60
13	10.00	1,523,627.60	85,545.65	1,581,890.14
14	10.00	1,581,890.14	88,112.02	1,643,155.93
15	10.00	1,643,155.93	90,755.38	1,707,640.60
16	10.00	1,707,640.60	93,478.04	1,775,578.81
17	10.00	1,775,578.81	96,282.39	1,847,226.07
18	10.00	1,847,226.07	99,170.86	1,922,860.73
19	10.00	1,922,860.73	102,145.98	2,002,786.22
20	10.00	2,002,786.22	105,210.36	2,087,333.45
21	10.00	2,087,333.45	108,366.67	2,176,863.45
22	10.00	2,176,863.45	111,617.67	2,271,770.35
23	10.00	2,271,770.35	114,966.20	2,372,484.56
24	10.00	2,372,484.56	118,415.19	2,479,476.31
25	10.00	2,479,476.31	121,967.65	2,593,259.53
26	10.00	2,593,259.53	125,626.68	2,714,396.14
27	10.00	2,714,396.14	129,395.48	2,843,500.73
28	10.00	2,843,500.73	133,277.34	2,981,245.73
29	10.00	2,981,245.73	137,275.66	3,128,367.08
30	10.00	3,128,367.08	141,393.93	3,285,670.46

Withdrawals Versus No Withdrawals Most financial planners understand the story of Henry. The important point of Henry's situation is that he took withdrawals from his portfolio during an unfortunate sequence of returns. During a period of regular withdrawals, it doesn't matter that his portfolio returns averaged 10 percent over the long run. It is the sequence of returns combined with regular withdrawals that was devastating to his portfolio. To

TABLE 7.5 The Actual Results

Year	Return %
1	–20.00
2	–10.00
3	9.00
4	8.00
5	12.00
6	–10.00
7	–2.00
8	25.00
9	27.00
10	61.00
Average Return	**10.00**

TABLE 7.6 Portfolio Balance Analysis

Year	Returns (%)	Withdrawal ($)	Ending Balance ($)
1	–20.00	60,000.00	752,000.00
2	–10.00	61,800.00	621,180.00
3	9.00	63,654.00	607,703.34
4	8.00	65,563.62	585,510.90
5	12.00	67,530.53	580,138.01
6	–10.00	69,556.44	459,523.41
7	–2.00	71,643.14	380,122.67
8	25.00	73,792.43	382,912.80
9	27.00	76,006.20	389,771.37
10	61.00	78,286.39	501,490.82

TABLE 7.7 Reversed Returns

Year	Return (%)	Withdrawal ($)	Ending Balance ($)
1	61.00	60,000.00	1,513,400.00
2	27.00	61,800.00	1,843,532.00
3	25.00	63,654.00	2,224,847.50
4	–2.00	65,563.62	2,116,098.20
5	–10.00	67,530.53	1,843,710.91
6	12.00	69,556.44	1,987,053.00
7	8.00	71,643.14	2,068,642.65
8	9.00	73,792.43	2,174,386.74
9	–10.00	76,006.20	1,888,542.48
10	–20.00	78,286.39	1,448,204.87

illustrate this point, imagine that Henry never took withdrawals from his portfolio (Table 7.8).

The time value of money comes into play when withdrawals are taken. When Henry experienced negative returns early in retirement while taking withdrawals, he had less money in his portfolio to grow over time. To maintain his inflation-adjusted withdrawal rate, Henry needed a bull market at the beginning of retirement.

TABLE 7.8 Returns Analysis Without Withdrawals

Actual Return Sequence with No Withdrawals

Year	Return (%)	Ending Balance ($)
1	–20.00	800,000.00
2	–10.00	720,000.00
3	9.00	784,800.00
4	8.00	847,584.00
5	12.00	949,294.08
6	–10.00	854,364.67
7	–2.00	837,277.38
8	25.00	1,046,596.72
9	27.00	1,329,177.84
10	61.00	2,139,976.32
Average Return	**10.00%**	

Reverse Return Sequence with No Withdrawals

Year	Return (%)	End Balance ($)
1	61.00	1,610,000.00
2	27.00	2,044,700.00
3	25.00	2,555,875.00
4	–2.00	2,504,757.50
5	–10.00	2,254,281.75
6	12.00	2,524,795.56
7	8.00	2,726,779.20
8	9.00	2,972,189.33
9	–10.00	2,674,970.40
10	–20.00	2,139,976.32
Average Return	**10.00%**	

Henry's retirement plan is deterministic because it assumes that returns will be the same each and every year. What Henry and Mr. D didn't understand was that averaging 10 percent over time is very different than getting 10 percent each and every year. As Henry left the office, Mr. D wished he had a more dynamic retirement planning process—one that allowed for varying variables.

Stochastic Planning Using Monte Carlo Simulation

Monte Carlo is a stochastic tool that helps people think in terms of probability and not certainty. As opposed to using a deterministic process, financial planners can use Monte Carlo to simulate risk in investment returns. A financial plan's probability of success can be tested by simulating the variability of investment returns. Typically, to measure this variability, the expected mean and standard deviation of the portfolio's investment returns are used in a Monte Carlo model. What would Mr. D have told Henry had this approach been used?

Using Henry's same information but an expected return of 10 percent with a standard deviation of 17.5 percent, Mr. D can assign success probabilities for how long Henry's money will last. Henry has a 64 percent chance that his portfolio will last 30 years (Figure 7.18). If Henry is not comfortable with that success rate, then Mr. D can increase both expected return and standard deviation, or decrease withdrawals. Mr. D could change the return to 20 percent, but this is obviously not realistic. In Henry's case, it makes more sense to decrease the withdrawal rate. Assuming that Henry will be comfortable with a 70 percent chance of success, then Mr. D needs to lower the annual withdrawal to $55,000 (Figure 7.19).

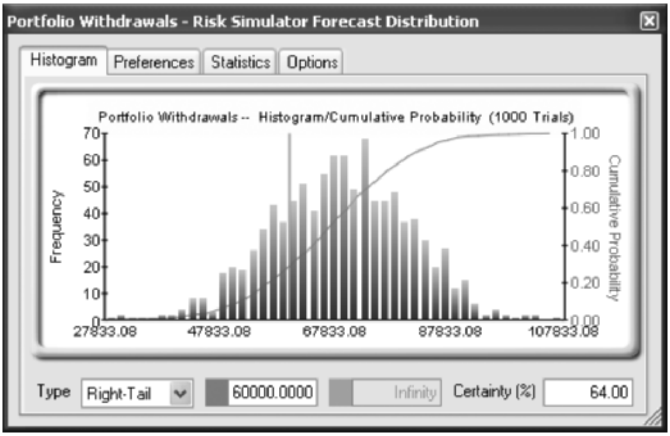


FIGURE 7.18 A 64 percent chance of portfolio survival at $60,000 withdrawals.

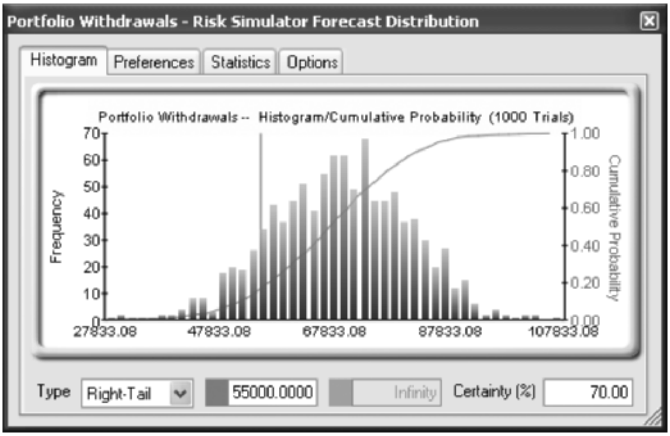


FIGURE 7.19 A 70 percent chance of portfolio survival at $55,000 withdrawals.

Expenses Lower Returns

It is truly a misuse of Monte Carlo and unfair to the client to illustrate a plan without fees if an advisory fee is to be charged. If Mr. Determinist charges Henry a 1 percent advisory fee, then this figure must be deducted from the annual return assumption, which will lower the plan's 30-year success probability to 54 percent. In Henry's case, the standard deviation will still be 17.5 percent, which is higher than a standard deviation of a portfolio that averages 9 percent. One can simply modify the Monte Carlo simulation to allow an advisory fee to be included by maintaining the return and standard deviation assumptions and deducting the advisory fee. For Henry's plan to still have a 70 percent success ratio after a 1 percent fee, he can withdraw an inflation-adjusted $47,000 annually, which is notably different from the $55,000 withdrawal rate before fees.

Success Probability

Monte Carlo educates the client about the trade-off between risk and return with respect to withdrawals. The risk is the success probability with which the client is comfortable. The return is the withdrawal rate. The financial planner should understand that a higher success rate amounts to lower withdrawals. A by-product of this understanding is that a higher success rate also increases the chance of leaving money in the portfolio at the client's death. In other words, Henry may be sacrificing lifestyle for an excessive probability of success. For Henry to have a 90 percent chance that his portfolio will

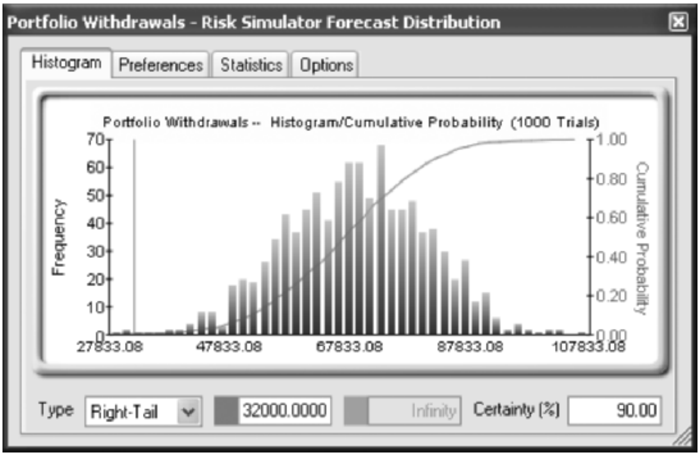


FIGURE 7.20 A 90 percent chance of portfolio survival at $32,000 withdrawals.

last 30 years, he needs to lower his withdrawals to $32,000 (Figure 7.20). An equally important interpretation of this result is that Henry has a 90 percent chance of dying with money in his portfolio. This is money he could have used for vacation, fancy dinners, gifts for his family, or circus tickets.

Success Tolerance

Going back to Henry's example of withdrawing $47,000 each year, if 5,000 simulation trials are run, a 70 percent success rate means that 3,500 times the plan worked. The 1,500 times the plan failed resulted in Henry being unable to take out $47,000 each and every year for 30 years. What is unclear about the 1,500 failures is how many of these resulted in a withdrawal amount marginally less than $47,000. If Henry takes out $47,000 for 29 years and then only withdraws $46,000 in the last year, is this a failure? Monte Carlo says yes. Most people are more flexible.

Establishing a success tolerance alleviates this problem. If Henry's goal is to take out $47,000 but he would be quite happy with $42,000, then he has a success tolerance of $5,000. This is the same as running a simulation using $42,000 with a zero success tolerance; however, the purpose of the success tolerance is to clearly illustrate to Henry the likelihood that a range of withdrawals will be achieved. By accounting for both the complexities of the market and the flexibility of human response to those complexities, Monte Carlo helps Henry understand, prepare for, and properly choose his risk tolerance.

Bear Markets and Monte Carlo

No matter what financial planning method is used, the reality is that a bear market early in retirement will drastically affect the plan. If Mr. D had used Monte Carlo when Henry first came to him and Henry took out $47,000 in Year 1 and $48,410 in Year 2, the portfolio balance at the end of the second year would have been $642,591. For the portfolio to last another 28 years and to preserve a 70 percent success rate, Henry must reduce his withdrawal amount to $31,500! The difficulty of this situation is obvious; however, Mr. D is in a position to help Henry make a decision about maintaining his standard of living versus increasing the chances of running out of money.

Table 7.9 illustrates running a Monte Carlo simulation at the end of each year to determine the withdrawal amount that preserves a 70 percent success rate for Henry's plan.

Like most people, Henry will not be enthusiastic about lowering his retirement salary by as much as 22 percent in any year. Without changing the return assumption, Henry's alternative is to accept a lower success rate. If Henry never adjusted his withdrawal rate from the initial $47,000, after 10 years his portfolio value would be $856,496 and his withdrawal would be $61,324 ($\$47,000 \times 1.03^9$). The success probability is 60 percent for a portfolio life of 20 years.

Other Monte Carlo Variables

Monte Carlo can simulate more than just investment returns. Other variables that are frequently simulated by financial planners using Monte Carlo include inflation and life expectancy.

TABLE 7.9 Simulation-Based Withdrawal Rates

Year	Return (%)	Beginning ($)	End Balance ($)	Monte Carlo Withdrawal ($)	Withdrawal Change (%)	Remaining Years
1	–20.00	1,000,000	762,400	47,000	0	29
2	–10.00	762,400	653,310	36,500	–22	28
3	9.00	653,310	676,683	32,500	–11	27
4	8.00	676,683	693,558	34,500	6	26
5	12.00	693,558	735,904	36,500	6	25
6	–10.00	735,904	627,214	39,000	7	24
7	–2.00	627,214	580,860	34,500	–12	23
8	25.00	580,860	685,137	32,750	–5	22
9	27.00	685,137	819,324	40,000	22	21
10	61.00	819,324	1,239,014	49,750	24	20

Inflation Since 1926, inflation has averaged approximately 3 percent annually with a standard deviation of 4.3 percent. In a plan with inflation-adjusted withdrawals, the change in inflation is significant. According to Ibbotson and Associates, inflation averaged 8.7 percent from the beginning of 1973 until the end of 1982. If such a period of inflation occurred at the beginning of retirement, the effect on a financial plan would be terrible.

Life Expectancy Using mortality tables, a financial planner can randomize the life expectancy of any client to provide a more realistic plan. According to the National Center for Health Statistics, the average American born in 2002 has a life expectancy of 77.3 years with a standard deviation of 10. However, financial planners should be more concerned with the specific probability that their clients will survive the duration of the plan.

Monte Carlo Suggestions

Financial plans created using Monte Carlo should not be placed on autopilot. As with most forecasting methods, Monte Carlo is not capable of simulating real-life adjustments that individuals make. As previously discussed, if a portfolio experienced severe negative returns early in retirement, the retiree can change the withdrawal amount. It is also important to realize that Monte Carlo plans are only as good as the input assumptions.

Distributions If Henry is invested in various asset classes, it is important for Mr. D to determine the distinct distribution characteristics of each asset class. The most effective approach to modeling these differences is by utilizing a distribution-fitting analysis in Risk Simulator.

Taxes Henry Tirement's situation involved a tax-deferred account and a pretax salary. For individuals with nontaxable accounts, rebalancing may cause taxes. In this case, a financial planner using Monte Carlo might employ a tax-adjusted return and a posttax salary might be used. The after-tax account balance should be used in the assumptions for clients with highly concentrated positions and a low tax basis who plan to diversify their investments.

Correlations It is important to consider any correlations between variables being modeled within Monte Carlo. Cross-correlations, serial correlations, or cross-serial correlations must be simulated for realistic results. For example, it may be shown that a correlation exists between investment returns and inflation. If this is true, then these variables should not be treated as independent of each other.

CASE STUDY: HOSPITAL RISK MANAGEMENT

This case is contributed by Lawrence Pixley, a founding partner of Stroudwater Associates, a management consulting firm for the health-care industry. Larry specializes in analyzing risk and uncertainty for hospitals and physician practices in the context of strategic planning and operational performance analyses. His expertise includes hospital facility planning, hospital/physician joint ventures, medical staff development, physician compensation packages utilizing a balanced scorecard approach, practice operations assessment, and practice valuations. Larry spent 15 years in health-care management, and has been a consultant for the past 23 years, specializing in demand forecasting using scientific management tools including real options analysis, Monte Carlo simulation, simulation-optimization, data envelopment analysis (DEA), queuing theory, and optimization theory. He can be reached at lpixley@stroudwaterassociates.com.

Hospitals today face a wide range of risk factors that can determine success or failure, including:

- Competitive responses both from other hospitals and physician groups.
- Changes in government rules and regulations.
- Razor-thin profit margins.
- Community relations as expressed through zoning and permitting resistance.
- State of the bond market and the cost of borrowing.
- Oligopsony (market with a few buyers) of a few large payers, for example, the state and federal governments.
- Success at fund-raising and generating community support.
- Dependence on key physicians, admitting preferences, and age of medical staff.
- High fixed cost structure.
- Advances in medical technology and their subsequent influence on admissions and lengths of stay.

In addition, hundreds of hospitals across the country are faced with aging facilities. Their dilemma is whether to renovate or relocate to a new site and build an entirely new facility. Many of these hospitals were first constructed in the early 1900s. Residential neighborhoods have grown up around them, locking them into a relatively small footprint, which severely hampers their options for expansion.

The Problem

Located in a large metropolitan area, CMC is a 425-bed community hospital. The region is highly competitive, with 12 other hospitals located within a 20-mile radius. Like most hospitals of similar size, CMC consists of a series of buildings constructed over a 50-year time span, with three major buildings 50, 30 and 15 years old. All three facilities house patients in double occupancy (or two-bed) rooms.

The hospital has been rapidly outgrowing its current facilities. In the last year alone, CMC had to divert 450 admissions to other hospitals, which meant a loss of $1.6 M in incremental revenue. Figure 7.21 shows CMC's average daily census and demonstrates why the hospital is running out of bed space.

Because of this growing capacity issue, the hospital CEO asked his planning team to project discharges for the next 10 years. The planning department performed a trend line analysis using the linear regression function in Excel and developed the chart shown in Figure 7.22. Applying a Poisson distribution to the projected 35,000 discharges, the planners projected a total bed need of 514. They made no adjustment for a change in the average length of stay over that 10-year period, assuming that it would remain constant. See Figure 7.23.

Confronted with the potential need to add 95 beds, the board of directors asked the CEO to prepare an initial feasibility study. To estimate the cost of adding 95 beds to the existing campus, the administrative staff first consulted with a local architect who had designed several small projects for the hospital. The architect estimated a cost of $260M to renovate the

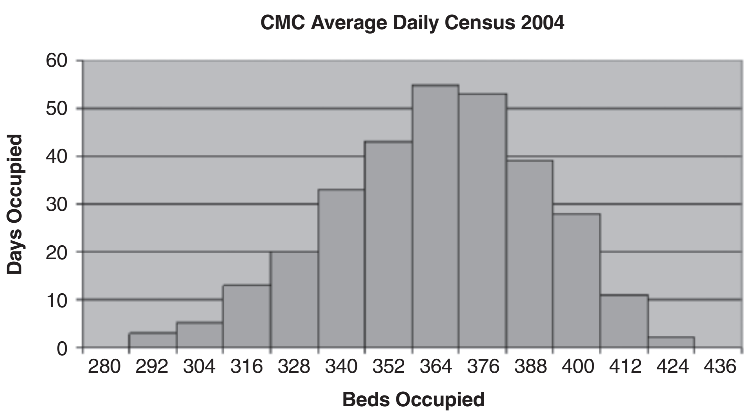


FIGURE 7.21 Histogram of CMC bed occupancy by number of days beds were occupied.

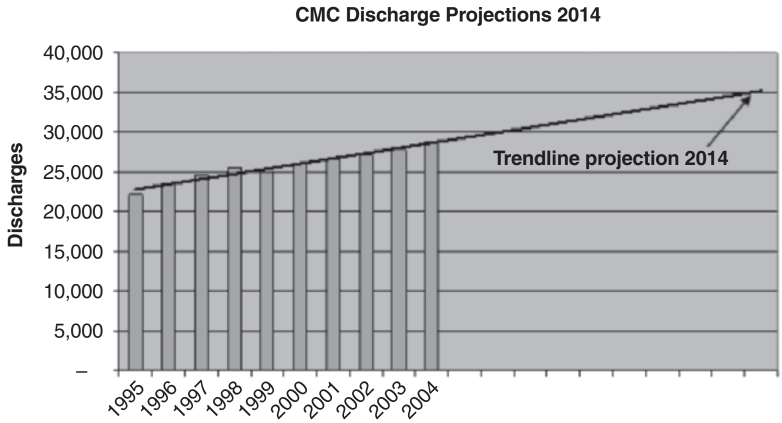


FIGURE 7.22 Trend line projections of CMC discharges for next 10 years (provided by CMC planning department).

existing structure and build a new addition, both of which were required to fit 95 more beds within the hospital's current footprint. To accommodate the additional beds on the current site, however, all beds would have to be double occupancy. Single occupancy rooms—the most marketable today—simply could not be accommodated on the present campus.

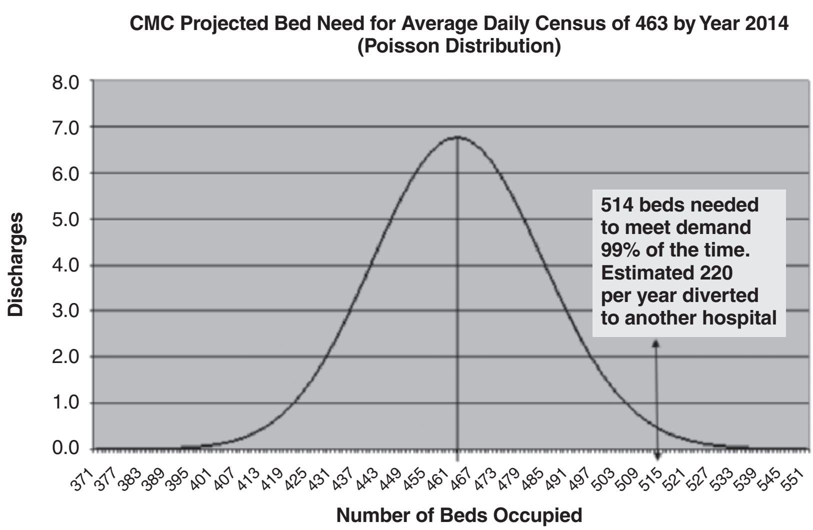


FIGURE 7.23 Projected CMC bed needs based on estimated average daily census of 463 patients for year 2014 (provided by CMC planning department).

In 1990, the hospital board faced a similar decision, whether to build a needed addition on the present campus or to relocate. The board opted to invest $90 million in a major expansion on the current site. Faced with the current dilemma, many of those same board members wished that in 1990 they had been able to better analyze their future options. A number of them expressed regrets that they did not relocate to another campus then. They clearly understood that their current decision—to renovate and add to the existing campus or to relocate—would be a decision the hospital would live with for the next 30 to 50 years.

There was no available site in the town (25 acres minimum), but there was space available in the adjacent town near a new $110 million ambulatory care center the hospital built five years ago. Yet, given the amount invested in the current campus and the uncertainty of how a new location would affect market share, there was real hesitancy to relocate.

The board had other considerations as well. Historically there had been litigation involved every time the hospital tried to expand. The neighboring property owners unsuccessfully opposed the Emergency Department expansion in 1999, but had managed through various legal actions to delay the construction three years. This delay added significantly to the cost of construction, in addition to the revenue lost from not having the modernized facility available as projected.

Two members of the board had attended a conference on the future of hospitals and noted that building more double occupancy rooms was not a good decision for the following reasons:

- By the time the facility was ready for construction, code requirements for new hospital construction would likely dictate single occupancy rooms.
- Patients prefer single rooms and CMC would be at a competitive disadvantage with other hospitals in the area that were already converting to single occupancy.
- Single occupancy rooms require fewer patient transfers and therefore fewer staff.
- Rates of infection were found to be considerably lower.

After receiving a preliminary cost estimate from the architect on a replacement hospital, the CFO presented the analysis shown in Figure 7.24 to the Finance Committee as an initial test of the project's viability. The initial projections for a new hospital estimated construction costs at $670 million. The study estimated a $50 million savings by not funding further capital improvements in the existing buildings. The CFO projected that the hospital would have a debt service capacity of an additional $95 million, assuming that the planning department's volume projections were accurate and that

Initial Capital Analysis for New Hospital ($ in M)		
Cost of Project	$	670
Less: Unrestricted Cash	$	(150)
: Deferred Maintenance	$	(50)
: Existing Debt Capacity	$	(100)
: Future Debt Capacity Based on New Volume	$	(95)
: Sale of Assets	$	(56)
: Capital Campaign	$	(150)
Capital Shortfall	$	69

FIGURE 7.24 Capital position analysis for new hospital as prepared by CMC chief financial officer.

revenue and expense per admission remained static. The balance would have to come from the sale of various properties owned by the hospital and a major capital campaign. Over the years, the hospital had acquired a number of outlying buildings for administrative functions and various clinics that could be consolidated into a new facility. In addition, there was a demand for additional residential property within the town limits, making the hospital's current site worth an estimated $17 million. Although skeptical, the CFO felt that with additional analysis, it could be possible to overcome the projected $69 million shortfall.

The board authorized the administration to seek proposals from architectural firms outside their area. The Selection Committee felt that given the risks of potentially building the wrong-sized facility in the wrong location, they needed firms that could better assess both risks and options. At the same time, as a hedge pending the completion of the analysis, the committee took a one-year option on the 25-acre property in the adjacent town. After a nationwide review, CMC awarded the project analysis to a nationally recognized architectural firm and Stroudwater Associates, with the strategic planning and analytics in Stroudwater's hands.

The Analysis

Stroudwater first needed to test the trend line projections completed by CMC's planning department. Rather than taking simple trend line projections based on past admissions, Stroudwater used a combination of both qualitative and quantitative forecasting methodologies. Before financial projections could be completed, a better estimate of actual bed need was required. Stroudwater segmented the bed need calculation into five key decision areas: population trends, utilization changes, market share, length of stay, and queuing decisions. Given the rapid changes in health-care technology in particular, it was determined that forecasting beyond 10 years was too speculative, and the board agreed that 10 years was an appropriate period for the analysis. In addition, the hospital wanted to project a minimum of 3 years beyond completion of hospital construction. Because projections were required for a minimum of 10 years, and because of the large number of variables involved, Stroudwater employed Monte Carlo simulation techniques in each of these five decision areas. See Figure 7.25.

For qualitative input to this process, the hospital formed a 15-person steering committee composed of medical staff, board directors, and key administrative staff. The committee met every three weeks during the four-month study and was regularly polled by Stroudwater on key decision areas through the entire process.

In addition, Stroudwater conducted 60 interviews with physicians, board members, and key administrative staff. During the interviews with key physicians in each major service line, Stroudwater consultants were struck by the number of aging physicians that were in solo practice and not planning to replace themselves, a significant risk factor for CMC. The CFO identified another issue: A majority of physicians in key specialties had recently stopped accepting insurance assignments, further putting the hospital at risk vis-à-vis its major competitor whose employed physicians accepted assignment from all payers.

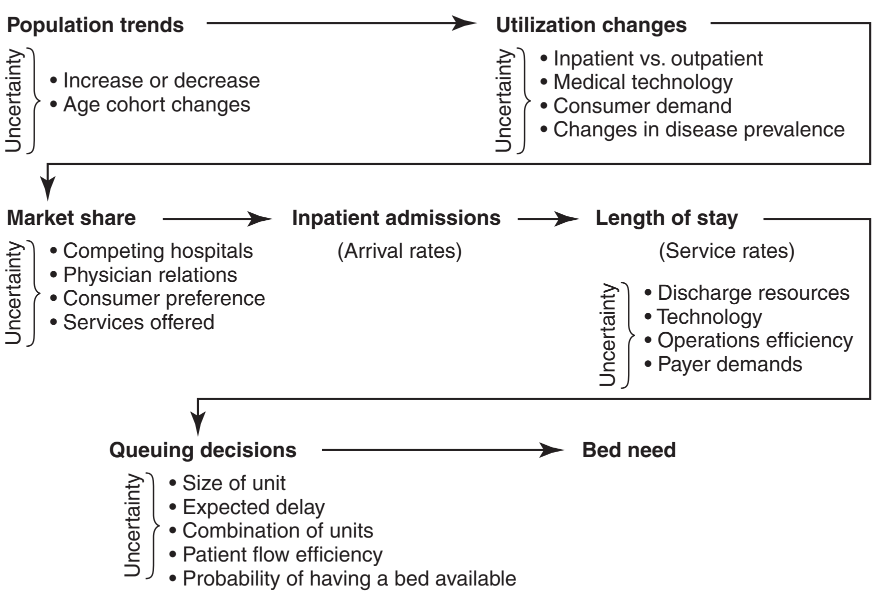


FIGURE 7.25 Stroudwater Associates methodology for forecasting hospital bed requirements.

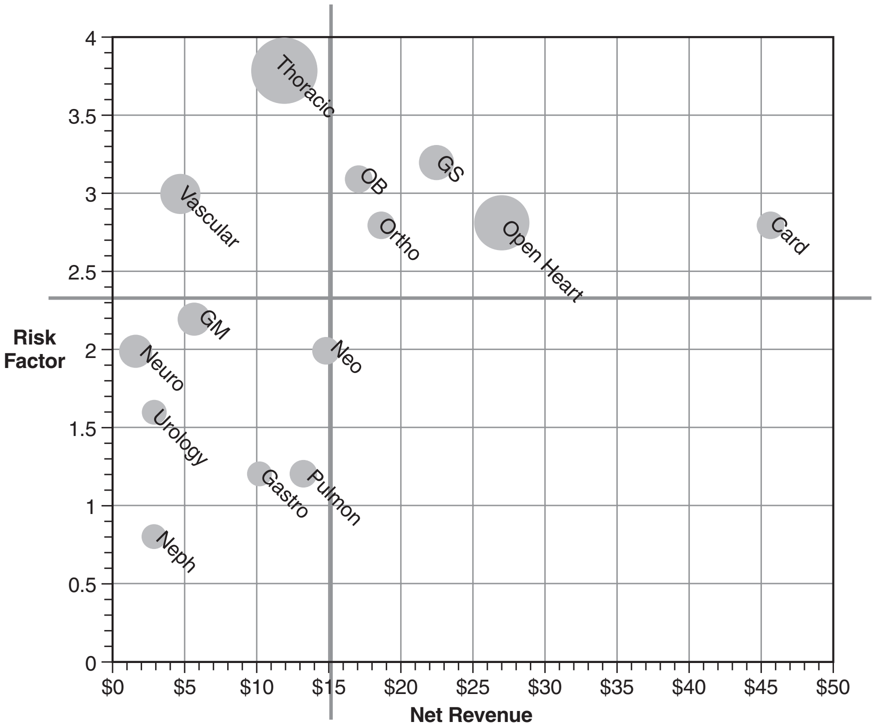


FIGURE 7.26 Bubble chart highlighting service lines considered most at risk (upper right quadrant). Operating margin is represented by the size of the bubble.

To understand better what service lines were at risk, Stroudwater developed a bubble diagram (Figure 7.26) to highlight areas that needed further business planning before making market share estimates. The three variables were net revenue, operating margin, and a subjective risk factor rating system.

The following risk factors were identified, assigned a weight, rated on a scale of one to five, and plotted on the y-axis:

- Size of practice—percentage of solo and two-physician practices in specialty.
- Average age of physicians in specialty.
- Potential competitive threat from other hospitals.
- Percentage of admissions coming from outside of service area.
- Percentage of physicians in the specialty accepting assignment from major insurance carriers.

The analysis revealed five key specialties—orthopedics, obstetrics, general surgery, open-heart surgery, and cardiology—in which CMC's bottom line was at risk, but which also afforded the greatest opportunity for future profitability. To better inform market share estimates, Stroudwater then developed mini business plans for each of the areas identified in the upper right-hand quadrant of Figure 7.26.

Population Trends To determine future population numbers in the CMC service area, Stroudwater depended on nationally recognized firms that specialize in population trending. Because hospital utilization is three times higher for over 65 populations, it was important to factor in the ongoing effect of the baby boomers. Stroudwater also asked members of the Steering Committee to review the 2014 population projections and determine what local issues not factored into the professional projections should be considered.

The committee members raised several concerns. There was a distinct possibility of a major furniture manufacturer moving its operations to China, taking some 3,000 jobs out of the primary service area. However, there was also the possibility of a new computer chip factory coming to the area. Stroudwater developed custom distributions to account for these population/employment contingencies.

Utilization Projections On completion of its population forecasting, Stroudwater turned its attention to calculating discharges per 1,000 people, an area of considerable uncertainty. To establish a baseline for future projections, 2004 discharge data from the state hospital association were used to calculate the hospitalization use rates (discharges per 1,000) for CMC's market. Stroudwater calculated use rates for 34 distinct service lines. See Table 7.10.

Stroudwater factored a number of market forces affecting hospital bed utilization into the utilization trend analyses. The consultants considered the following key factors that might decrease facility utilization:

- Better understanding of the risk factors for disease, and increased prevention initiatives (e.g., smoking prevention programs, cholesterol-lowering drugs).
- Discovery/implementation of treatments that cure or eliminate diseases.
- Consensus documents or guidelines that recommend decreases in utilization.
- Shifts to other sites causing declines in utilization in the original sites
  - As technology allows shifts (e.g., ambulatory surgery).
  - As alternative sites of care become available (e.g., assisted living).

TABLE 7.10 Utilization Trends for 2014 by Service Line

Product Line	2004 Discharges	2004 Length of Stay	2004 Population	2004 Discharges 1000	2004 Days 1000	2004 Average Length of Stay	2014 Population	2014 Change in Utilization (%)	2014 Estimated Total Market Discharges
Abortion	137	213	1,193,436	0.12	0.18	1.6	1,247,832	0	1
Adverse Effects	878	2,836	1,193,436	0.74	2.40	3.2	1,247,832	0	9
AIDS and Related	358	3,549	1,193,436	0.30	3.00	9.9	1,247,832	0	3
Burns	66	859	1,193,436	0.07	0.73	10.0	1,247,832	0	
Cardiology	19,113	75,857	1,193,436	16.17	64.19	4.0	1,247,832	18	20,1
Dermatology	435	3,446	1,193,436	0.37	2.92	7.9	1,247,832	0	4
Endocrinology	3,515	18,246	1,193,436	2.97	15.44	5.2	1,247,832	5	3,7
Gastroenterology	9,564	46,103	1,193,436	8.09	39.01	4.8	1,247,832	5	10,0
General Surgery	7,488	51,153	1,193,436	6.34	43.28	6.8	1,247,832	9	7,9
Gynecology	3,056	6,633	1,193,436	2.59	7.31	2.6	1,247,832		3,2
Hematology	1,362	10,325	1,193,436	1.15	8.74	7.6	1,247,832	8	1,4
Infectious Disease	2,043	15,250	1,193,436	1.73	12.90	7.5	1,247,832	0	2,1
Neonatology	1,721	20,239	1,193,436	1.46	17.13	11.8	1,247,832	4	1,8
Neurology	5,338	34,873	1,193,436	4.52	29.51	6.5	1,247,832	12	5,8
Neurosurgery	3,042	13,526	1,193,436	2.57	11.45	4.4	1,247,832	12	3,2
Newborn	11,197	25,007	1,193,436	9.47	21.16	2.2	1,247,832	–5	11,6
Obstetrics	13,720	36,962	1,193,436	11.61	31.28	2.7	1,247,832	–5	14,4
Oncology	1,767	11,563	1,193,436	1.50	9.76	6.5	1,247,832	15	1,5

Source: State Hospital Discharge Survey.

- Changes in practice patterns (e.g., encouraging self-care and healthy lifestyles, reduced length of hospital stay).
- Changes in technology.

Stroudwater also considered the following factors that may increase hospital bed utilization:

- Growing elderly population.
- New procedures and technologies (e.g., hip replacement, stent insertion, MRI).
- Consensus documents or guidelines that recommend increases in utilization.
- New disease entities (e.g., HIV/AIDS, bioterrorism).
- Increased health insurance coverage.
- Changes in consumer preferences and demand (e.g., bariatric surgery, hip and knee replacements).

In all key high-volume services, Stroudwater consultants made adjustments for utilization changes and inserted them into the spreadsheet model, using a combination of uniform, triangular, and normal distributions.

Market Share The Steering Committee asked Stroudwater to model two separate scenarios, one for renovations and an addition to the current campus, and the second for an entirely new campus in the adjacent town. To project the number of discharges that CMC was likely to experience in the year 2014, market share assumptions for both scenarios were made for each major service line.

A standard market share analysis aggregates zip codes into primary and secondary service markets depending on market share percentage. Instead, Stroudwater divided the service area into six separate market clusters using market share, geographic features, and historic travel patterns.

Stroudwater selected eight major service areas that represented 80 percent of the admissions for further analysis and asked committee members and key physicians in each specialty area to project market share. The committee members and participating physicians attended one large meeting where CMC planning department members and Stroudwater consultants jointly presented results from the mini-business plans. Local market trends and results of past patient preference surveys were considered in a discussion that followed. As an outcome from the meeting, participants agreed to focus on specific factors to assist them in estimating market share, including:

- Change in patient preference.
- Proximity of competing hospitals.
- New hospital “halo” effect.

- Change in "hospital of choice" preferences by local physicians.
- Ability to recruit and retain physicians.

Using a customized survey instrument, Stroudwater provided those participating in the exercise with four years of trended market share data; challenging them to create a worst-case, most likely, and best-case estimate for (1) each of the six market clusters in (2) each of the eight service lines for (3) each campus scenario.

After compiling the results of the survey instrument, Stroudwater assigned triangular distributions to each variable. An exception to the process occurred in the area of cardiac surgery. There was considerable discussion over the impact of a competing hospital potentially opening a cardiothoracic surgery unit in CMC's secondary service market. For the "current campus" scenario, the Steering Committee agreed that if a competing unit were opened it would decrease their market share to the 15 to 19 percent range, and they assigned a 20 percent probability that their competitor would open the unit. Should the competitor not build the unit, a minority of the group felt that CMC's market share would increase significantly to the 27 to 30 percent range; a 30 percent probability was assigned. The remaining members were more conservative and estimated a 23 to 25 percent market share. Similarly, estimates were made for the new campus in which participants felt there were better market opportunities and where losses would be better mitigated should the competing hospital open a new cardiothoracic unit.

Stroudwater used the custom distributions shown in Figure 7.27.

Average Length of Stay Stroudwater performed length of stay estimates for 400 diagnostic groupings (DRG) using a combination of historic statistics from the National Hospital Discharge Survey of the National Center for Health Statistics and actual CMC data.

Key CMC physicians participated in estimating length of stay based on the benchmark data, their knowledge of their respective fields, and historic CMC data. Stroudwater consultants separately trended historic lengths of stay and developed an algorithm for weighting benchmark data and CMC physician estimates. Length of stay estimates were rolled up into one distribution for each of the major service lines.

At this point, Stroudwater performed a sensitivity analysis (Figure 7.28) to determine which assumptions were driving the forecasts. Based on the relative unimportance population had on outcome, the population distribution assumptions were dropped in favor of single point estimates.

Queuing Decisions A typical approach to determining bed need, and the one used by the CMC Planning Department, is to multiply projections for single point admissions by those for single point lengths of stay to determine

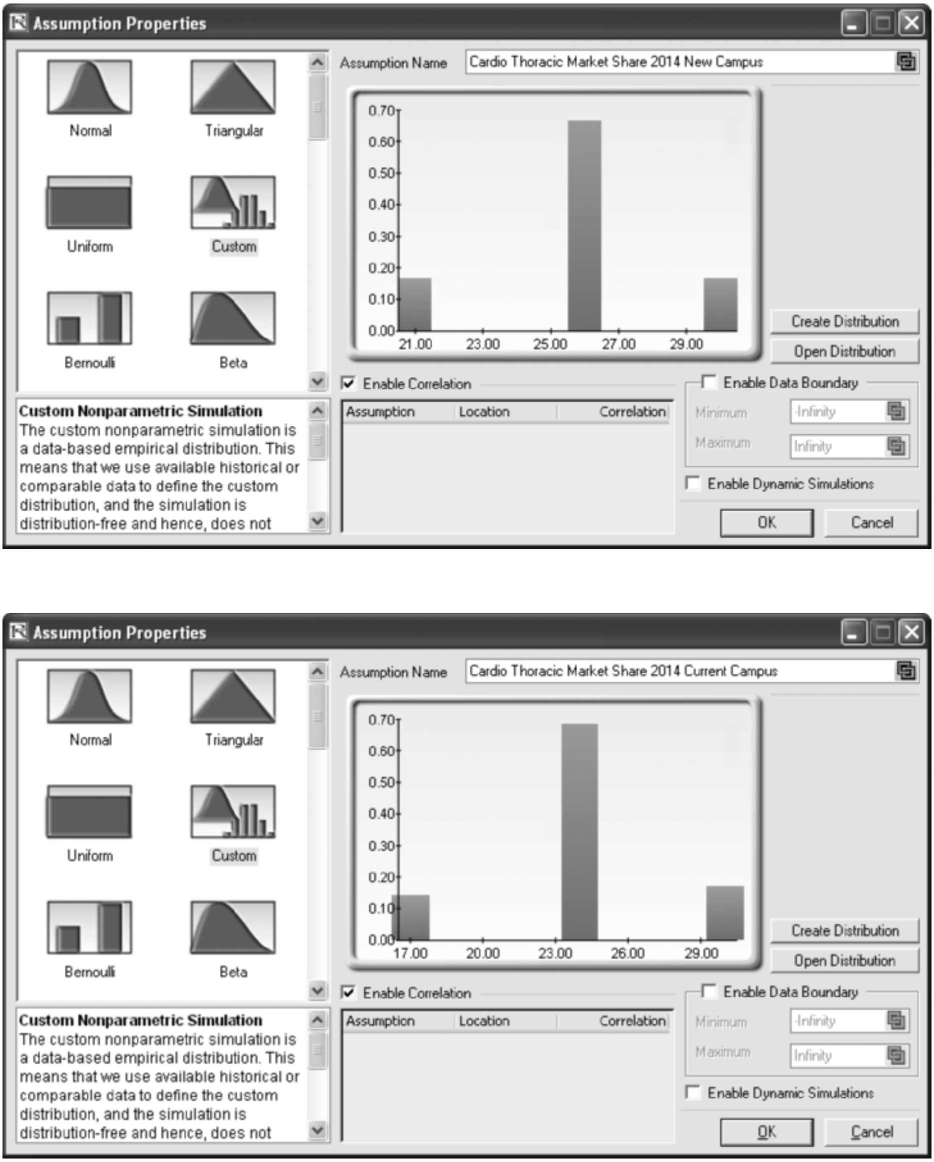


FIGURE 7.27 Cardiothoracic market share using custom distributions comparing market share assumptions for both current and new campus.

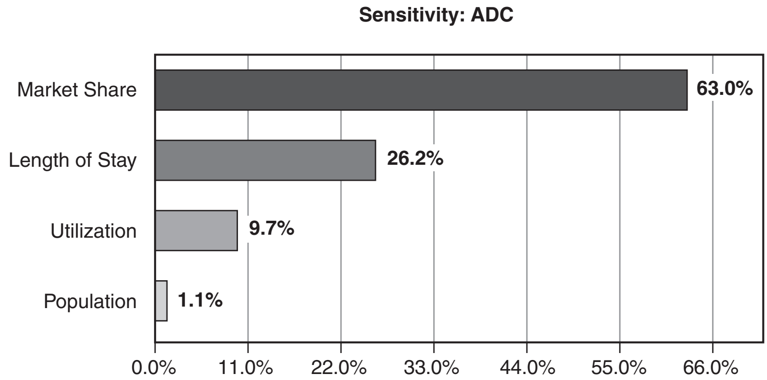


FIGURE 7.28 Sensitivity analysis of key variables in Monte Carlo simulation.

the total number of patient days. Patient days are divided by 365 to determine the average daily census (ADC). A Poisson distribution is then applied to the ADC to determine the total number of beds required. In addition to the problems of single point estimates, Poisson distributions assume that all arrivals are unscheduled and thus overstate the bed need if any of the services have elective or urgent admissions.

Because CMC had categorized all of its admissions by urgency of the need for a bed, Stroudwater was able to conduct an analysis for each unit and found wide differences in the timing needs for beds ranging from OB with 100 percent emergency factor to Orthopedics with 57 percent of its admissions classified as elective. See Table 7.11.

To deepen the analysis, the physician members of the committee met separately to determine which units could be combined because of natural affinities and similar nursing requirements. The Steering Committee then met to discuss service targets for each category of admission. They agreed

TABLE 7.11 Orthopedic/Neurosurgery Admissions Classified by Admission Priority

	Emergency	Urgent	Elective	Total
Total Days	5,540	415	7,894	13,849
Total Admissions	1,497	112	2,133	3,743
Percentage (Admits)	40%	3%	57%	100%

TABLE 7.12 MGK Blocking Model Showing Bed Need Service Targets

				Bed Needs Service Target		
Unit	Discharges Arrival Rates	Service Rate 1/ALOS	CV	Emergency < 1 day	Urgent 1–2 days	Elective 2–3 days
Medical Cardiology	8.6301	0.0606	142.3973	71%	25%	4%
General Surgery	10.9315	0.0741	147.5753	49%	2%	49%
Orthopedics	17.9795	0.0901	199.5719	40%	3%	57%

that "Emergencies" had to have a bed available immediately, "Urgent" within 48 hours, and "Elective" within 72 hours. Using a multiple channel queuing model jointly developed by Dr. Johnathan Mun and Lawrence Pixley, bed needs were determined for each of the major unit groupings. See Table 7.12 and Table 7.13.

Distributions had been set for utilization and market share by service line to determine the arrival rates needed for the queuing model. Length of stay distributions by service line had been determined for the service rate input to the model. Forecast cells for Monte Carlo simulation were set for "Probability of Being Served" for <1, 1–2, and 2–3 days for each of the units respectively.

As its planning criteria, the committee decided on a target rate of 95 percent confidence in having a bed available with a greater than 50 percent certainty. Stroudwater employed an iterative process to the model, rerunning the Monte Carlo simulation until the performance criteria were met. For example, the first run for Orthopedics at 75 beds had a certainty of 47.8 percent at 95 percent confidence level compared to a later run of 78 beds with a certainty of 60.57 percent. The 78-bed figure was adopted. See Figure 7.29.

The Results of the Analysis

The committee's perception was that a new hospital located in a neighboring community closer to its target markets would improve market share in key specialties. That perception was reinforced by Stroudwater's findings in the projected differences in bed need between the two sites. See Table 7.14.

The project architects utilized the bed demand information and completed construction cost projections for each of the two scenarios. With a need for only 39 additional beds on the current campus compared to the

TABLE 7.13 MGK Blocking Model with Determination of Beds and Probability of Availability

Unit	Period/Day 3 No. Beds Per Day	No. Beds Per Period	Beds Busy	Prob. Busy	Prob. Served < 1 Day	Prob. Served 1–2 Days	Prob. Served 2–3 Days
Medical Cardiology	102	34	34	76.3%	99.4%	100.0%	100.0%
General Surgery	66	22	22	84.7%	89.6%	100.0%	100.0%
Orthopedics/ Neuro	78	26	26	81.9%	96.0%	100.0%	100.0%
Total		82					

original projection of the need for 95 additional beds, the architects were able to design space that afforded 92 private rooms.

The architects estimated the project cost for the new replacement facility at $587 million compared to $285 million for the renovation/addition option for the current campus. The new campus solution afforded an estimated increase in capital campaign contributions of $125 million and income from sale of assets of $56 million, bringing the borrowing required to an estimated $231 million. Borrowing for the current campus option was estimated to be $110 million.

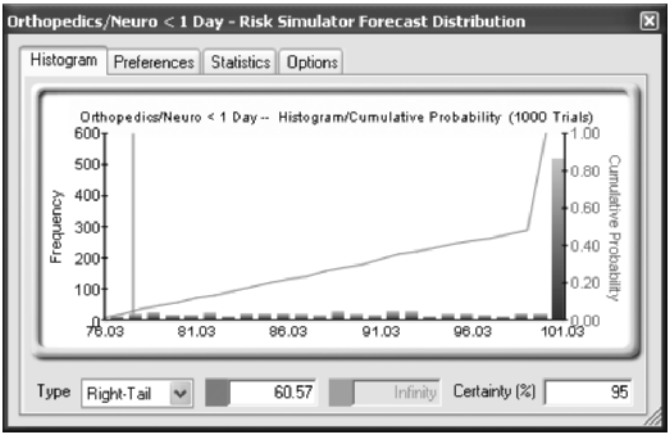


FIGURE 7.29 Frequency distribution for 78 orthopedic beds at new campus site.

TABLE 7.14 Results of Bed Need Projections for Both Current and New Campus Solutions

Service Line	2004 Current Campus	2014 Projections Current Campus	2014 Projections New Campus
Obstetrics	47	48	49
Cardiology	41	43	47
Pulmonary	50	55	56
Infectious Disease	18	20	19
Ortho/Neurosurgery	49	69	73
Rehabilitation	16	18	18
Hematology/Oncology	14	15	16
General Surgery	38	41	42
Vascular/Cardiac Surg.	64	60	68
Urology	14	14	16
Gastroenterology	18	21	21
Neurology	18	20	21
Other Medical	12	14	15
Other Surgical	25	26	28
Total Beds	425	464	489

The pro formas reflected the following advantages to the new campus solution:

- Revenue per admission and per bed was higher with the new campus scenario because of the expected increase in higher margin specialty admissions. Cardiothoracic surgery, for example, contributed $11,600 per case in margin compared to $2,200 for Urology.
- CMC was averaging 6.1 full-time equivalent (FTE) employees per bed in the current facility, much of it due to facility inefficiencies. Stroudwater projected that a renovated campus could bring down the FTE to occupied bed ratio to 6.0 but projected the new facility at 5.8.
- Utility costs were projected to drop from the current $4.51 to $4.08 per square foot and maintenance costs were expected to drop from $2.46 to $1.40 per square foot.
- Loss of revenue from disruption of operations would be minimized with the new campus solution.
- The adjacent community provided assurances to CMC that it would not experience zoning difficulties should the hospital choose to relocate, whereas because of ongoing community opposition to further construction on the existing campus, a three-year delay in construction was expected.

In addition to the foregoing pro forma presentations (see Tables 7.15 and 7.16), Stroudwater provided the board with the Monte Carlo simulation results for projected profit margin in the year 2014 as shown in Figure 7.30. Interestingly, the profit margins projected for the two scenarios were remarkably similar, with the new hospital scenario having a slightly higher probability of exceeding a 4 percent profit margin. Given the similar outcomes of the pro formas, the board elected to proceed with the new campus solution. They felt that even though moving to the adjacent community was a risk, the risk of remaining on the current site was even greater. They realized that their future expansion options were limited should the projections prove to underrepresent future demand for services, whereas the new campus afforded them a great deal of flexibility for unanticipated events.

A bond rating agency rewarded CMC's approach to risk assessment with a favorable rating. Its opinion letter reflected the following observations:

- CMC received high marks for the decision-making process. The agency appreciated the alternative analysis of building on the present campus compared to a new campus and the unique approach of incorporating uncertainty into the calculation of bed need. It noted that the original projections for a 515-bed facility were scaled back to 489 beds as a result of the analysis.
- CMC received points for involving the physicians in the Steering Committee, and for the fact that CMC administration continually met with the medical staff to provide updates on the analysis.
- The agency felt that the relocation to the new campus was a risk by moving away from existing physician offices, but the risk was not only mitigated but enhanced by a privately owned and developed 300,000 square foot medical office building as part of the new campus. (It noted the lack of room for medical office facilities on the existing campus.) It also accepted the argument that CMC's long-term financial viability was improved by the future ability to recruit and retain physicians, particularly in large group practices.
- The fact that the new hospital would be located adjacent to CMC's ambulatory care center that had already been in full operation for 6 years was also viewed positively as patients were accustomed to traveling to this site.
- The agency found that management had compellingly examined all reasonable scenarios for patient volume and third-party reimbursement and their impact on earnings and liquidity.

The following were the principal advantages of using applied risk analysis in this case:

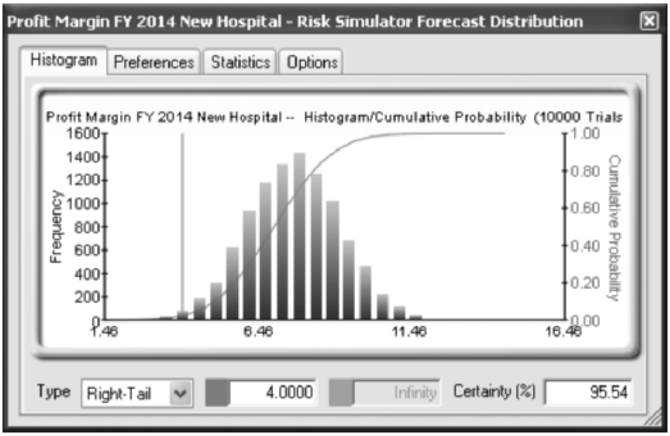
- Board members, many of whom were familiar with applied risk analysis in their own industries, were more comfortable making a major relocation

TABLE 7.15 Pro Forma for New Hospital Scenario

	FY2008	FY2007	FY2008	FY2010	FY2011	FY2012
Total Operating Revenue	$338,250,000	$350,550,000	$358,360,000	$364,000,000	$361,088,000	$382,720,000
Total Expenses	$314,215,000	$325,641,000	$332,003,200	$336,336,000	$320,762,624	$328,852,160
EBIDA	$ 31,613,900	$ 32,487,900	$ 33,935,700	$ 35,242,900	$ 44,325,376	$ 57,867,840
EBIDA Margin	9.2%	9.1%	9.3%	9.5%	12.2%	15.1%
Total Capital and Other Costs	$ 10,602,167	$ 10,774,367	$ 10,883,707	$ 10,962,667	$ 34,891,167	$ 36,583,630
Operating Income /(Loss)	$ 13,432,833	$ 14,134,633	$ 15,473,093	$ 16,701,333	$ 5,434,209	$ 17,284,210
Operating Margin	4.0%	4.0%	4.3%	4.6%	1.5%	4.5%
Contributions and Investment Income	$ 7,578,900	$ 7,578,900	$ 7,578,900	$ 7,578,900	$ 4,000,000	$ 4,000,000
Net Income/(Loss)	$ 21,011,733	$ 21,713,533	$ 23,051,993	$ 24,280,233	$ 9,434,209	$ 21,284,210
Profit Margin	6.1%	6.1%	6.3%	6.5%	2.6%	5.5%
Income Available for Capital	$ 31,613,900	$ 32,487,900	$ 33,935,700	$ 35,242,900	$ 44,325,376	$ 57,867,840
Debt Service Coverage Ratio	6.1	6.1	6.2	3.3	1.3	3.5

TABLE 7.16 Pro Forma for Current Campus Scenario

	FY2008	FY2007	FY2008	FY2010	FY2011	FY2012
Total Operating Revenue	$338,250,000	$350,550,000	$358,360,000	$364,000,000	$361,088,000	$370,760,000
Total Expenses	$314,215,000	$325,641,000	$332,003,200	$336,336,000	$321,845,888	$326,759,160
EBIDA	$ 31,613,900	$ 32,487,900	$ 33,935,700	$ 35,242,900	$ 43,242,112	$ 48,000,840
EBIDA Margin	9.2%	9.1%	9.3%	9.5%	11.9%	12.9%
Total Capital and Other Costs	$ 10,602,167	$ 10,774,367	$ 10,883,707	$ 10,962,667	$ 23,318,486	$ 23,370,578
Operating Income/(Loss)	$ 13,432,833	$ 14,134,633	$ 15,473,093	$ 16,701,333	$ 15,923,626	$ 20,630,262
Operating Margin	4.0%	4.0%	4.3%	4.6%	4.4%	5.6%
Contributions and Investment Income	$ 7,578,900	$ 7,578,900	$ 7,578,900	$ 7,578,900	$ 4,000,000	$ 4,000,000
Net Income/(Loss)	$ 21,011,733	$ 21,713,533	$ 23,051,993	$ 24,280,233	$ 19,923,626	$ 24,630,262
Profit Margin	6.1%	6.1%	6.3%	6.5%	5.5%	6.6%
Income Available for Capital	$ 31,613,900	$ 32,487,900	$ 33,935,700	$ 35,242,900	$ 43,242,112	$ 48,000,840
Debt Service Coverage Ratio	6.1	6.1	6.2	3.3	2.7	6.1



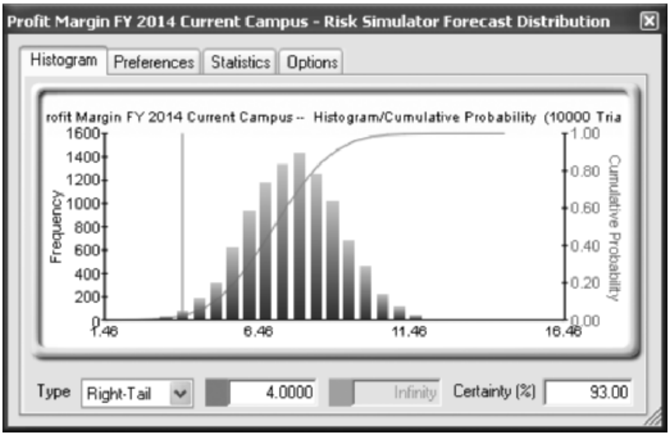


FIGURE 7.30 Frequency distribution of profit margin comparing alternative scenarios.

decision based on a range of probable outcomes rather than on previously employed single point estimates.

- The bond-rating agency awarded the hospital a favorable bond rating because "what if" scenarios were employed and because of the methods utilized in both identifying and mitigating risk factors.
- The hospital was able to reduce the number of projected beds and hence its overall construction cost because of the more sophisticated queuing methodology employed.

CASE STUDY: RISK-BASED EXECUTIVE COMPENSATION VALUATION

This case was written by Patrick Haggerty, a principal at the executive compensation consulting firm James F. Reda & Associates, LLC. As independent advisors to management and boards, the firm assists companies with designing and implementing executive compensation programs. The firm has significant expertise in valuing long-term incentive awards using guidance provided by FASB Statement No. 123 (Revised 2004), Share-Based Payment (FAS 123(R)), and related interpretations. Through partnering with Dr. Mun and using his option valuation software packages, James F. Reda & Associates, LLC, helps clients determine and understand the compensation expense impact of selecting alternative long-term incentive designs.

This case is based on actual projects performed, but for the purposes of maintaining proprietary information, we use a fictitious entity named Boris Manufacturing, Inc. (Boris). This case study is about the process that Boris used to evaluate alternative long-term incentive (LTI) plan designs and determine the fair value for expensing purposes, as required by the new financial accounting standards. Through the following steps, the management team and the compensation committee worked together to evaluate the advantages and disadvantages of the various LTI vehicles available. The steps undertaken included:

- Reviewing the historical LTI awards made to employees.
- Reviewing the company's LTI plan.
- Conducting a market study.
- Evaluating advantages and disadvantages of each LTI vehicle available.

Ultimately, Boris decided to award restricted stocks that vest on achieving a total shareholder return target. Because the performance condition is total shareholder return, an option-pricing model can be used to determine fair value based on a barrier option, where the stock vests only after breaching a predetermined upper performance barrier. A simple Black–Scholes is not designed to value these types of awards. Instead, Monte Carlo and binomial lattice models like Dr. Mun's Real Options Super Lattice Solver and Risk Simulator software are most appropriate because they include the necessary input factors. FAS 123(R) considers the vesting criteria on Boris's restricted stock award a "market condition," meaning it is stock-price related. This distinction is important because if Boris designed a plan that vests on achieving a non-stock price-related measure (i.e., earnings per share, or EPS, and earnings before interest, taxes, depreciation, and amortization, or EBITDA), the company could not factor the performance condition into the fair value of the award (FAS 123(R) calls this type of performance measure a "performance condition"). For more technical details on valuing regular employee stock options based on the 2004 revised FAS 123, see Dr. Johnathan Mun's case study in Chapter 14 on valuing employee stock options.

Background

Boris Manufacturing Inc. is a publicly traded billion-dollar manufacturer of-chemical products. The company has 2,000 employees with approximately 200 management- and executive-level employees. The compensation committee at Boris is responsible for determining executive pay levels and awarding LTIs to all employees. The compensation committee evaluated pay practices among its peer group companies and determined that LTIs should be a significant and important part of total compensation. Accordingly, the company has awarded its management- and executive-level employees LTIs. Historically, Boris awarded stock options to employees because prior to FAS 123(R) the expense was zero—under previous accounting rules, compensation expense was zero for at-the-money stock options if the number of shares awarded are known on the grant date.

Boris's stock option awards have not provided the incentive or link to shareholders that the compensation committee expected. Over the past 4 years, Boris's stock price has been relatively volatile and has generally decreased. Roughly half of the stock options Boris awarded to employees have an exercise price higher than the current stock price or being *underwater*. Further, the company kept awarding more stock options because the stock price continued to fall. As a result, the company has unproductive stock overhang, employees with minimal linkage to shareholders, and few shares remaining in their stock pool. As described next, the compensation committee decided to undertake a study to evaluate these issues.

Compensation Committee Process

To review alternative LTI designs, the compensation committee conducted the following steps:

1. Reviewed historical LTI awards made to employees.

   Purpose: To understand what employees had received in the past such as type of award, current fair value of award, and any gains received.

   Result: Over the past 3 years, Boris awarded approximately 900,000 stock options to employees each year (2.7 million in total). Unfortunately, roughly half are underwater, and very few employees were able to exercise and sell with any gain.

2. Reviewed company's long-term incentive plan.

   Purpose: To understand types of LTI vehicles that Boris shareholders approved in its LTI plan and how many shares are available for awards.

   Result: Boris's LTI plan is very flexible and allows for all types of LTI vehicles, including:

   - Nonqualified stock options (NQSO).
   - Incentive stock options (ISO).
   - Stock-settled stock appreciations rights (Stock SAR).
   - Restricted stock and restricted stock units (RSU).
   - Performance shares and performance units.

   Due to higher than expected stock option grants made over the past 3 years, the company has only 500,000 shares available for future grants. It is likely that Boris will need to go back to shareholders next year, so they want to use the remaining shares wisely.

3. Conducted a market study.

   Purpose: To determine competitive practices for LTI awards, costs, and LTI designs (vesting, performance measures, termination provisions, and holding periods).

   Result: Based on an analysis of industry competitors, the company determined that historical stock option awards were above market levels—on an individual position level, overhang basis, and cost basis. Also, it was determined that many peer group companies are awarding full value shares (i.e., restricted stock and performance shares) rather than stock options. Among the peer group companies that are awarding full value shares with performance conditions, the most common performance conditions were total shareholder return, earnings per share, and EBITDA.

4. Evaluated advantages and disadvantages of each LTI vehicle available. Table 7.17 summarizes the compensation committee's findings.

Compensation Committee Decision

The compensation committee decided to award restricted stock that vests on achieving a predefined total shareholder return (TSR) target. Key factors that influenced the committee to select this LTI plan included:

- Reduction in overhang and run rate.
- Better link to shareholders.
- Requires minimum acceptable level of performance before payout.
- Promotes stock ownership because executives do not have to sell shares to exercise.

TABLE 7.17 Advantages and Disadvantages of LTI Vehicles

LTI Vehicle	FAS 123R Measurement Approach	Key Employee Tax Issue	Key Advantage	Key Disadvantage
NQSOs	Fixed: grant date fair value[a]	Ordinary income tax at exercise	Determine taxable event, upside potential	Potential underwater, highly dilutive
ISOs	Fixed: grand date fair value	Capital gains tax at sale[b]	Capital gains, upside potential	No company tax deduction, ISO rules
Stock SARs	Fixed: grant date fair value	Ordinary income tax at exercise	Limits dilution, upside potential	Potential underwater
Restricted Stock	Fixed: grant date face value[c]	Ordinary income when vested	Retention, no cost to employee	Pay tax when vested, not 162(m) qualified
Restricted Stock Units (paid in stock)	Fixed: grant date face value	Ordinary income when delivered	Flexibility, can include performance	Flexibility subject to 409A rules
Performance Shares	Fixed/variable: stock price fixed, shares adjusted[d]	Ordinary income when vested	Additional shares and higher stock price	Setting performance measures
Performance Units (paid in cash)	Variable:[e] adjusted until paid	Ordinary income when vested	Receive cash, diversify	Cash flow, variable accounting

[a]Fair value based on an option-pricing model, such as Black–Scholes.

[b]If requisite holding periods are met, otherwise same as NQSOs.

[c]Face value equals stock price on grant date.

[d]Stock price fixed on grant date; shares are variable until measurement period is complete.

[e]Mark-to-market accounting until award is paid.

Details of the design include:

Type	Restricted stock
Vesting criteria	Vests on achieving 6 percent annual TSR
Performance period	3 years (average cumulative TSR must exceed 6 percent)
Dividend rights	Participants do not receive dividends until stock has vested
Number of shares	All-or-nothing award, no adjustment in number of shares if TSR is below or above 6 percent

Before selecting the 6 percent TSR target, the compensation committee reviewed Boris's historical TSR. Based on this review, it was determined that Boris's 3-year historical average annualized return is 5.2 percent, and using this and the volatility estimates, we were able to compute the expected distribution of future returns (see Figure 7.31). The committee considered this and set the TSR target and expected range TSR performance at:

TSR Target:	6%
Minimum Expected:	0%
Most Likely:	5%
Max Expected:	9%

The compensation committee considered and analyzed but ultimately decided against the following alternative plan designs. Each one of these alternatives would result in a different fair value calculation.

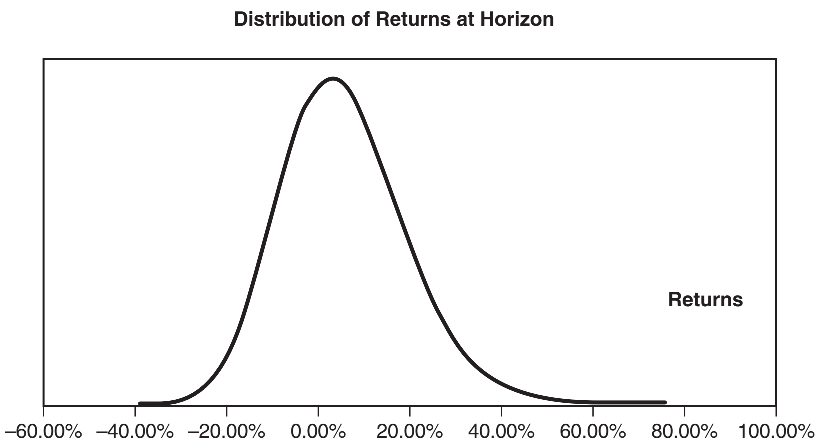


FIGURE 7.31 Boris's projected returns based on historical performance.

- Increasing duration of performance period from 3 years to 5 years.
- Vesting award based on company TSR performance against a peer group rather than a predetermined target.
- Awarding performance shares rather than restricted stock (Note: this change does not impact the fair value but impacts the number of shares that will vest).

Compensation Cost Determination

Using FAS 123(R) guidance, Boris determined the fair value of the restricted stock award for expense recognition. Compensation cost for the award will equal the fair value multiplied by the number of restricted shares granted. Determining the fair value for its restricted stock awards is similar to the process Boris had used to determine fair value of its stock option awards under the pro forma disclosure rules of FAS 123. However, a simple Black–Scholes model cannot be used to determine the fair value of an award with a TSR target. Instead, a Monte Carlo simulation model coupled with a binomial lattice model must be used with inputs as detailed next (see Chapters 12 and 13 for details on option valuation techniques). A Monte Carlo simulation model coupled with a binomial model is more appropriate than other closed-form option-pricing models because this analysis has a barrier associated with the payoff structure (i.e., TSR targets), which means only a binomial lattice can be used to model such barrier options. In addition, the potential that Boris's TSR will exceed these targets is highly uncertain and thus we need to run a Monte Carlo simulation to capture its expected value. Therefore, we couple Risk Simulator's Monte Carlo simulation capabilities with the Employee Stock Options Valuation and Real Options SLS software to perform the computations. See the chapters on real options analysis for more details on running the SLS software, or refer to the author's *Real Options Analysis, Second Edition* (Wiley Finance, 2005). The following are the assumptions used in the model:

- *Grant date.* This assumption determines the grant date stock price and interest rate assumption.
- *Grant date stock price.* Equals the closing stock price on the grant date, or $20.00 for this example.
- *Purchase price.* Typically $0 for restricted stock awards.
- *Volatility.* Calculated based on historical stock prices, 30 percent for this example. Significant guidance for determining this assumption is provided in FAS 123(R) and SEC's Staff Accounting Bulletin No. 107.
- *Contractual period.* Equals duration of performance period, 3 years for this example.
- *Dividend yield.* Calculated based on Boris's historical dividend yield, 1 percent for this example.

Dividend Payment Date	Dividend Amount	Stock Price ($)	Quarterly Dividend Yield (%)
3/15/2005	$0.04	15.00	0.27
6/15/2005	$0.04	15.50	0.26
9/15/2005	$0.04	15.75	0.25
12/15/2005	$0.04	16.00	0.25
Sum of quarterly dividend yields			1.03

- *Interest rate.* Based on the U.S. Treasury rates available on the grant date with a maturity equaling the contractual term. For this example, we used a 4 percent interest rate.
- *TSR target.* Boris's compensation committee set the target at 6 percent based on the company's 3-year historical average annualized return of 5.2 percent.
- *Expected range TSR performance.* Sets the parameters for determining the likelihood of achieving the TSR target. The committee thought it would be reasonable to assume a minimum expected TSR of 0 percent and a stretch TSR of 9 percent.
- *Suboptimal exercise multiple.* Set the price at which the participant is expected to exercise. This assumption is set at 10,000, which theoretically renders it unattainable. If this award were a stock option, this assumption could be used if employee exercise behavior indicated a lower level.

The results generated using Risk Simulator's Monte Carlo simulation coupled with the Real Options SLS provides a fair value of $10.27 (Figure 7.32). Real Options SLS software was used to obtain the restricted stock's fair-market valuation while Risk Simulator was used to simulate the potential TSR values. Thus, if Boris awards 400,000 restricted shares to employees, the compensation cost equals 400,000 × $10.27 = $4,108,000, which is accrued over the performance period of 3 years. If the Monte Carlo simulation model were not used, Boris would be required to use the grant date stock price, $20, resulting in an expense of 400,000 × $20 = $8,000,000. Therefore, by applying the right methodologies as well as the right engineered LTI grants, Boris was able to reduce its expenses by almost 50 percent.

Conclusion

Monte Carlo Simulation models can be used to help design the LTI award by understanding the impact that certain changes have on fair value, and to determine the fair value of the LTI award for expense purposes under FAS 123R. Without the use of such sophisticated methodologies, the fair value

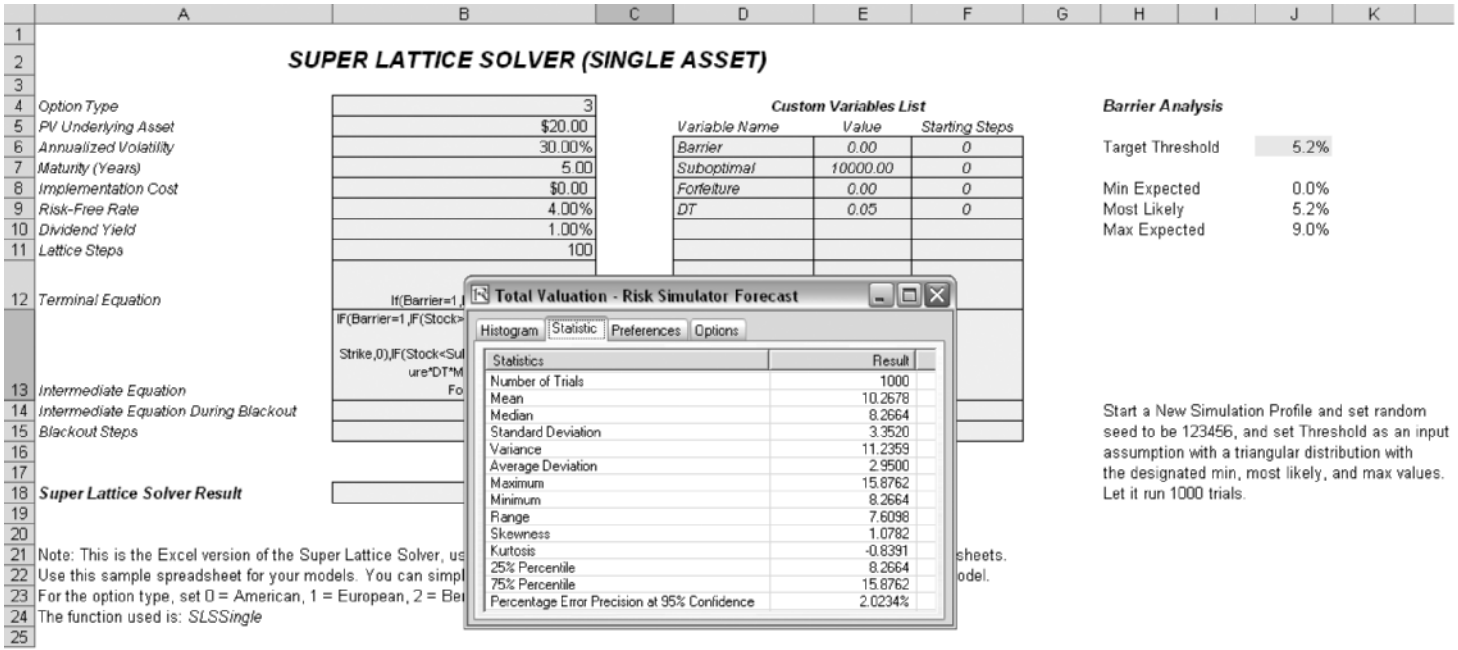


FIGURE 7.32 Total valuation results (sample only) for Boris's LTI.

would never have been computed correctly and the decision to undertake the right LTI would have been flawed. In addition, such methodologies outlined here can also be used for multiple other applications such as engineering LTIs and stock-based compensations that are tied to, say, a market index such as the S&P 500, or a company's performance (i.e., we can use financial metrics such as net profit margin, gross profits, EBITDA, and the like), or perhaps to some commodity price (e.g., price of gold or oil). For technical and application details on FAS 123R and running the Employee Stock Options Valuation software, please refer to the author's book, *Valuing Employee Stock Options (Under 2004 FAS 123)* (Wiley Finance, 2004).

PART

Five

Risk Prediction

CHAPTER 8

Tomorrow's Forecast Today

Forecasting is the act of predicting the future, whether it is based on historical data or speculation about the future when no history exists. When historical data exist, a quantitative or statistical approach is best, but if no historical data exist, then a qualitative or judgmental approach is usually the only recourse. Figure 8.1 lists the most common methodologies for forecasting.

DIFFERENT TYPES OF FORECASTING TECHNIQUES

Generally, forecasting can be divided into quantitative and qualitative approaches. Qualitative forecasting is used when little to no reliable historical, contemporaneous, or comparable data exists. Several qualitative methods exist such as the Delphi or expert opinion approach (a consensus-building forecast by field experts, marketing experts, or internal staff members), management assumptions (target growth rates set by senior management), as well as market research or external data or polling and surveys (data obtained through third-party sources, industry and sector indexes, or from active market research). These estimates can be either single-point estimates (an average consensus) or a set of prediction values (a distribution of predictions). The latter can be entered into Risk Simulator as a custom distribution and the resulting predictions can be simulated; that is, running a nonparametric simulation using the prediction data points as the custom distribution.

For quantitative forecasting, the available data or data that needs to be forecasted can be divided into time-series (values that have a time element to them, such as revenues at different years, inflation rates, interest rates, market share, failure rates, and so forth), cross-sectional (values that are time-independent, such as the grade point average of sophomore students across the nation in a particular year, given each student's levels of SAT scores, IQ, and number of alcoholic beverages consumed per week), or mixed panel (mixture between time-series and panel data, e.g., predicting sales over the next

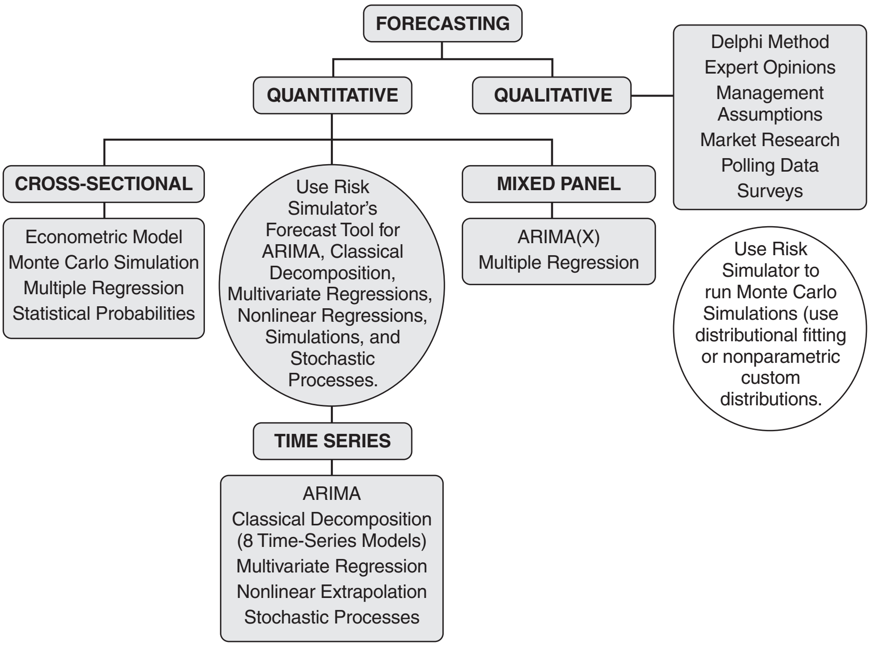


FIGURE 8.1 Forecasting methods.

10 years given budgeted marketing expenses and market share projections, which means that the sales data is time-series but exogenous variables such as marketing expenses and market share exist to help to model the forecast predictions).

The Risk Simulator software provides the user several forecasting methodologies:

- Time-Series Analysis
- Multivariate Regression
- Stochastic Forecasting
- Nonlinear Extrapolation
- Box–Jenkins ARIMA

RUNNING THE FORECASTING TOOL IN RISK SIMULATOR

In general, to create forecasts, several quick steps are required:

1. Start Excel and enter in or open existing historical data.

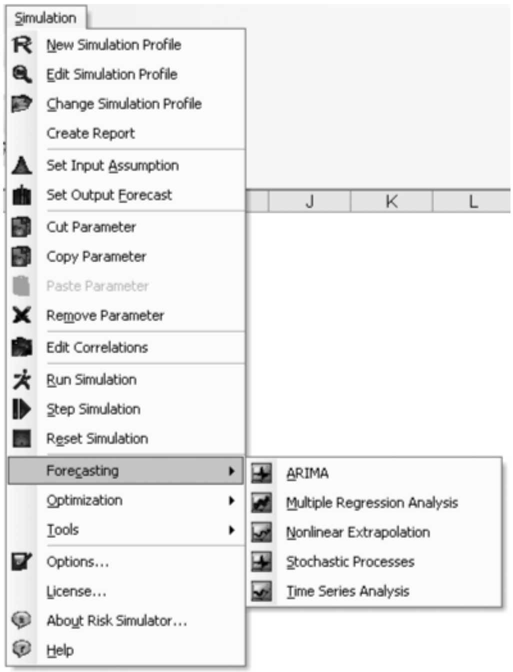


FIGURE 8.2 Risk Simulator's forecasting methods.

2. Select the data and click on *Simulation* | *Forecasting*.
3. Select the relevant sections (Box–Jenkins ARIMA, Time-series Analysis, Multivariate Regression, Stochastic Forecasting, or Nonlinear Extrapolation) and enter the relevant inputs.

Figure 8.2 illustrates the *Forecasting* tool and the various methodologies available in Risk Simulator.

The following provides a quick review of each methodology and several quick getting started examples in using the software. The example data files used to create these examples are included in the Risk Simulator software and can be accessed through: *Start* | *Programs* | *Real Options Valuation* | *Risk Simulator* | *Examples*.

TIME-SERIES ANALYSIS

Theory

Figure 8.3 lists the eight most common time-series models, segregated by seasonality and trend. For instance, if the data variable has no trend or

FIGURE 8.3 The eight most common time-series methods.

seasonality, then a single moving-average model or a single exponential-smoothing model would suffice. However, if seasonality exists but no discernible trend is present, either a seasonal additive or seasonal multiplicative model would be better, and so forth.

Procedure

Follow the steps listed to run a time-series analysis:

1. Start Excel and type in or open an existing spreadsheet with the relevant historical data (the following example uses the *Time-Series Forecasting* file in the examples folder).
2. Select the historical data, not including the variable name (data should be listed in a single column).
3. Select *Simulation* | *Forecasting* | *Time-Series Analysis*.
4. Choose the model to apply, enter the relevant assumptions, and click OK.

Make sure you start a new simulation profile or that there is an existing profile in the model if you want the forecast results to automatically generate simulation assumptions.

To follow along in this example, choose Auto Model Selection, enter 4 for seasonality periods per cycle, and forecast for 4 periods. See Figure 8.4.

Historical Sales Revenues

Year	Quarter	Period	Sales
2000	1	1	$684.20
2000	2	2	$584.10
2000	3	3	$765.40
2000	4	4	$892.30
2001	1	5	$885.40
2001	2	6	$677.00
2001	3	7	$1,006.60
2001	4	8	$1,122.10
2002	1	9	$1,163.40
2002	2	10	$993.20
2002	3	11	$1,312.50
2002	4	12	$1,545.30
2003	1	13	$1,596.20
2003	2	14	$1,260.40
2003	3	15	$1,735.20
2003	4	16	$2,029.70
2004	1	17	$2,107.80
2004	2	18	$1,650.30
2004	3	19	$2,304.40
2004	4	20	$2,639.40

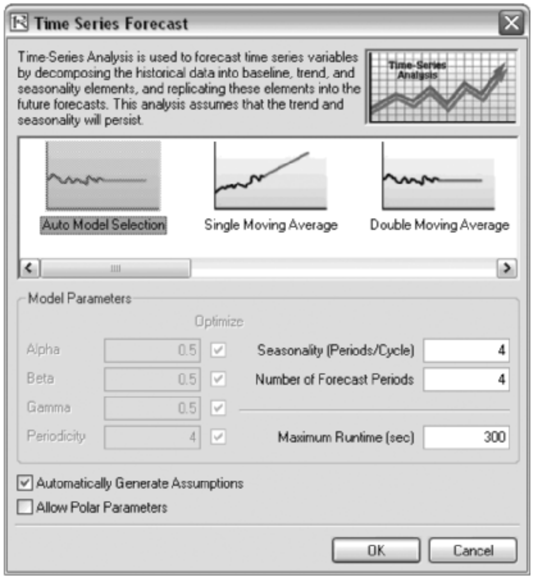


FIGURE 8.4 Time-series analysis.

Results Interpretation

Figure 8.5 illustrates the sample results generated by using the *Forecasting* tool. The model used was a Holt–Winters multiplicative model. Notice that in Figure 8.5, the model-fitting and forecast chart indicate that the trend and seasonality are picked up nicely by the Holt–Winters multiplicative model. The time-series analysis report provides the relevant optimized alpha, beta, and gamma parameters, the error measurements, fitted data, forecast values, and forecast-fitted graph. The parameters are simply for reference. Alpha captures the memory effect of the base level changes over time, beta is the trend parameter that measures the strength of the trend, while gamma measures the seasonality strength of the historical data. The analysis decomposes the historical data into these three elements and then recomposes them to forecast the future. The fitted data illustrates the historical data as well as the fitted data using the recomposed model and shows how close the forecasts are in the past (a technique called *backcasting*). The forecast values are either single-point estimates or assumptions (if the automatically generated assumptions option is chosen and if a simulation profile exists). The graph

Holt-Winters' Multiplicative

Summary Statistics

Alpha, Beta, Gamma	RMSE	Alpha, Beta, Gamma	RMSE
0.00, 0.00, 0.00	914.824	0.00, 0.00, 0.00	914.824
0.10, 0.10, 0.10	415.322	0.10, 0.10, 0.10	415.322
0.20, 0.20, 0.20	187.202	0.20, 0.20, 0.20	187.202
0.30, 0.30, 0.30	118.795	0.30, 0.30, 0.30	118.795
0.40, 0.40, 0.40	101.794	0.40, 0.40, 0.40	101.794
0.50, 0.50, 0.50	102.143		

The analysis was run with alpha = 0.2429, beta = 1.0000, gamma = 0.7797, and seasonality = 4

Time-Series Analysis Summary

When both seasonality and trend exist, more advanced models are required to decompose the data into their base elements: a base-case level (L) weighted by the alpha parameter, a trend component (b) weighted by the beta parameter, and a seasonality component (S) weighted by the gamma parameter. Several methods exist but the two most common are the Holt-Winters' additive seasonality and Holt-Winters' multiplicative seasonality methods. In the Holt-Winters' additive model, the base case level, seasonality, and trend are added together to obtain the forecast fit.

The best-fitting test for the moving average forecast uses the root mean squared errors (RMSE). The RMSE calculates the square root of the average squared deviations of the fitted values versus the actual data points.

Mean Squared Error (MSE) is an absolute error measure that squares the errors (the difference between the actual historical data and the forecast-fitted data predicted by the model) to keep the positive and negative errors from cancelling each other out. This measure also tends to exaggerate large errors by weighting the large errors more heavily than smaller errors by squaring them, which can help when comparing different time-series models. Root Mean Square Error (RMSE) is the square root of MSE and is the most popular error measure, also know as the quadratic loss function. RMSE can be defined as the average of the absolute values of the forecast errors and is highly appropriate when the cost of the forecast errors is proportional to the absolute size of the forecast error. The RMSE is used as the selection criteria for the best-fitting time-series model.

Mean Absolute Percentage Error (MAPE) is a relative error statistic measured as an average percent error of the historical data points and is most appropriate when the cost of the forecast error is more closely related to the percentage error than the numerical size of the error. Finally, an associated measure is the Theil's U statistic, which measures the naivety of the model's forecast. That is, if the Theil's U statistic is less than 1.0, then the forecast method used provides an estimate that is statistically better than guessing.

Period	Actual	Forecast Fit
1	684.20	
2	584.10	
3	765.40	
4	892.30	
5	885.40	684.20
6	677.00	667.55
7	1006.60	935.45
8	1122.10	1198.09
9	1163.40	1112.48
10	993.20	887.95
11	1312.50	1348.38
12	1545.30	1546.53
13	1596.20	1572.44
14	1260.40	1299.20
15	1735.20	1704.77
16	2029.70	1976.23
17	2107.80	2026.01
18	1650.30	1637.26
19	2304.40	2245.93
20	2639.40	2643.09
Forecast 21		2713.69
Forecast 22		2114.79
Forecast 23		2900.42
Forecast 24		3293.81

Error Measurements	
RMSE	71.8132
MSE	5157.1348
MAD	53.4071
MAPE	4.50%
Theil's U	0.3054

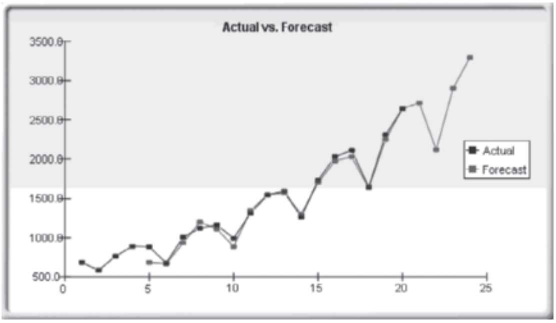


FIGURE 8.5 Example Holt–Winters forecast report.

illustrates the historical, fitted, and forecast values. The chart is a powerful communication and visual tool to see how good the forecast model is.

Notes

This time-series analysis module contains the eight time-series models seen in Figure 8.3. You can choose the specific model to run based on the trend and seasonality criteria or choose the Auto Model Selection, which will automatically iterate through all eight methods, optimize the parameters, and find the best-fitting model for your data. Alternatively, if you choose one of the eight models, you can also deselect the *optimize* checkboxes and enter your own alpha, beta, and gamma parameters (Figure 8.4). In addition, you would need to enter the relevant seasonality periods if you choose the automatic model selection or any of the seasonal models. The seasonality input has to be a positive integer (e.g., if the data is quarterly, enter 4 as the number of seasons or cycles a year, or enter 12 if monthly data, or any other positive integer representing the data periods of a full cycle). Next, enter the number of periods to forecast. This value also must be a positive integer. The maximum runtime is set at 300 seconds. Typically, no changes are required. However, when forecasting with a significant amount of historical data, the analysis might take slightly longer, and if the processing time exceeds this runtime, the process will be terminated. You can also elect to have the forecast automatically generate assumptions; that is, instead of single-point estimates, the forecasts will be assumptions. However, to automatically generate assumptions, a simulation profile must first exist. Finally, the polar parameters option allows you to optimize the alpha, beta, and gamma parameters to include zero and one. Certain forecasting software allows these polar parameters while others do not. Risk Simulator allows you to choose which to use. Typically, there is no need to use polar parameters. See Chapter 9 for the technical details on time-series forecasting using the eight decomposition methods.

MULTIVARIATE REGRESSION

Theory

It is assumed that the user is sufficiently knowledgeable about the fundamentals of regression analysis. The general bivariate linear regression equation takes the form of $Y = \beta_0 + \beta_1 X + \varepsilon$ where β_0 is the intercept, β_1 is the slope, and ε is the error term. It is bivariate as there are only two variables, a Y or dependent variable, and an X or independent variable, where X is also known as the regressor (sometimes a bivariate regression is also known as a univariate regression as there is only a single independent variable X). The

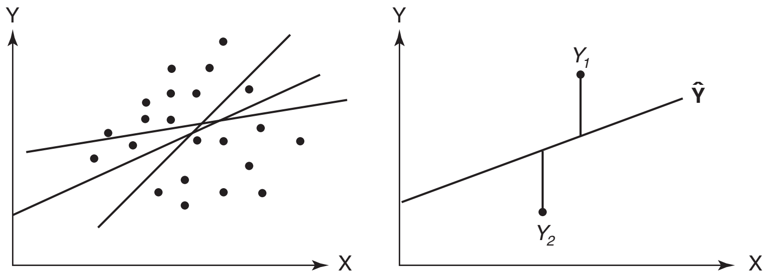


FIGURE 8.6 Bivariate regression.

dependent variable is named as such as it *depends* on the independent variable, for example, sales revenue depends on the amount of marketing costs expended on a product's advertising and promotion, making the dependent variable sales and the independent variable marketing costs. An example of a bivariate regression is seen as simply inserting the best-fitting line through a set of data points in a two-dimensional plane as seen on the left panel in Figure 8.6. In other cases, a multivariate regression can be performed, where there are multiple or k number of independent X variables or regressors, where the general regression equation will now take the form of $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \ldots + \beta_k X_k + \varepsilon$. In this case, the best-fitting line will be within a $k + 1$ dimensional plane.

However, fitting a line through a set of data points in a scatter plot as in Figure 8.6 may result in numerous possible lines. The best-fitting line is defined as the single unique line that minimizes the total vertical errors, that is, the sum of the absolute distances between the actual data points (Y_i) and the estimated line ($\hat{Y}$) as shown on the right panel of Figure 8.6. To find the best-fitting unique line that minimizes the errors, a more sophisticated approach is applied, using regression analysis. Regression analysis therefore finds the unique best-fitting line by requiring that the total errors be minimized, or by calculating

$$Min \sum_{i=1}^{n} \left(Y_i - \hat{Y}_i \right)^2$$

where only one unique line minimizes this sum of squared errors. The errors (vertical distances between the actual data and the predicted line) are squared to avoid the negative errors from canceling out the positive errors. Solving this minimization problem with respect to the slope and intercept requires calculating first derivatives and setting them equal to zero:

$$\frac{d}{d\beta_0}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2 = 0 \text{ and } \frac{d}{d\beta_1}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2$$

which yields the bivariate regression's least squares equations:

$$\beta_1 = \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right)}{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2} = \frac{\sum_{i=1}^{n} X_i Y_i - \frac{\sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i}{n}}{\sum_{i=1}^{n} X_i^2 - \frac{\left(\sum_{i=1}^{n} X_i\right)^2}{n}}$$

$$\beta_0 = \bar{Y} - \beta_1 \bar{X}$$

For multivariate regression, the analogy is expanded to account for multiple independent variables, where $Y_i = \beta_1 + \beta_2 X_{2,i} + \beta_3 X_{3,i} + \varepsilon_i$ and the estimated slopes can be calculated by:

$$\hat{\beta}_2 = \frac{\sum Y_i X_{2,i} \sum X_{3,i}^2 - \sum Y_i X_{3,i} \sum X_{2,i} X_{3,i}}{\sum X_{2,i}^2 \sum X_{3,i}^2 - \left(\sum X_{2,i} X_{3,i}\right)^2}$$

$$\hat{\beta}_3 = \frac{\sum Y_i X_{3,i} \sum X_{2,i}^2 - \sum Y_i X_{2,i} \sum X_{2,i} X_{3,i}}{\sum X_{2,i}^2 \sum X_{3,i}^2 - \left(\sum X_{2,i} X_{3,i}\right)^2}$$

In running multivariate regressions, great care must be taken to set up and interpret the results. For instance, a good understanding of econometric modeling is required (e.g., identifying regression pitfalls such as structural breaks, multicollinearity, heteroskedasticity, autocorrelation, specification tests, nonlinearities, and so forth) before a proper model can be constructed.

Procedure

Use the following steps to run a multivariate regression:

1. Start Excel and type in or open your existing data set (the illustration below uses the file *Multiple Regression* in the examples folder).
2. Check to make sure that the data is arranged in columns, select the data including the variable names, and click on *Simulation* | *Forecasting* | *Multiple Regression*.

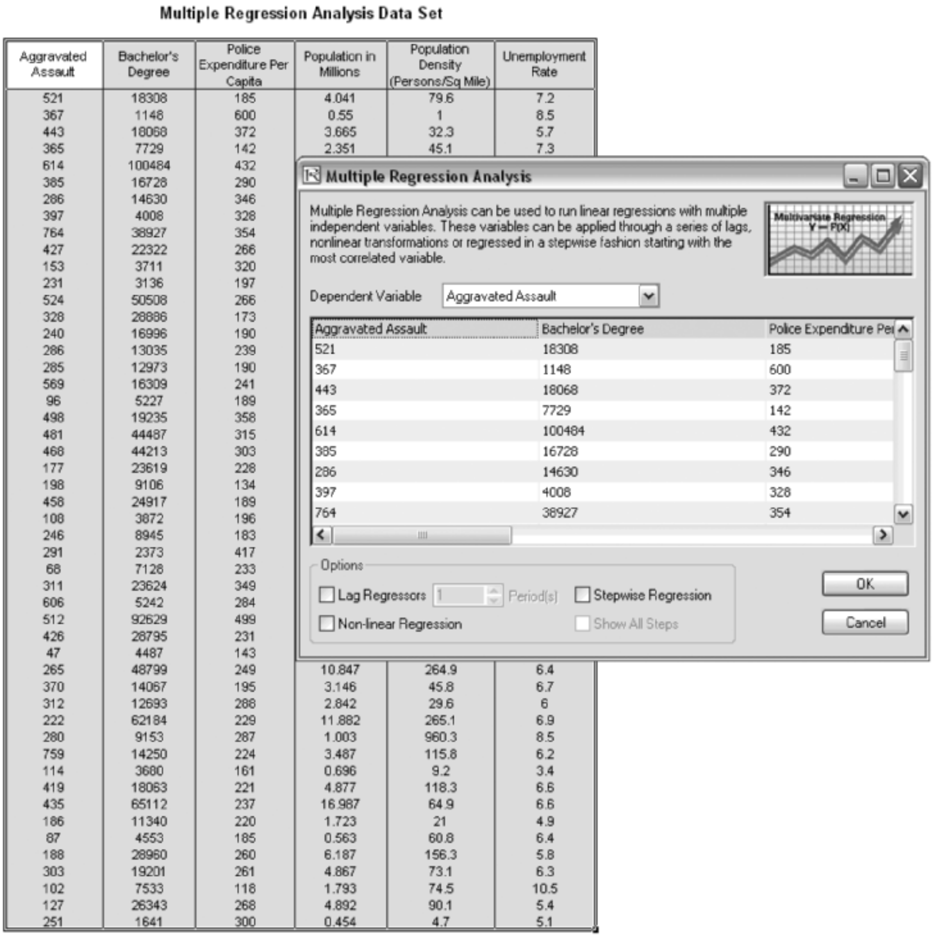


FIGURE 8.7 Running a multivariate regression.

3. Select the dependent variable and check the relevant options (lags, stepwise regression, nonlinear regression, and so forth), and click OK (Figure 8.7).

Results Interpretation

Figure 8.8 illustrates a sample multivariate regression result report generated. The report comes complete with all the regression results, analysis of variance results, fitted chart, and hypothesis test results. See Chapter 9 for the technical details on interpreting the results from a regression analysis.

STOCHASTIC FORECASTING

Theory

A stochastic process is nothing but a mathematically defined equation that can create a series of outcomes over time, outcomes that are not deterministic in nature; that is, an equation or process that does not follow any simple discernible rule such as price will increase *X* percent every year or revenues will increase by this factor of *X* plus *Y* percent. A stochastic process is by definition nondeterministic, and one can plug numbers into a stochastic process equation and obtain different results every time. For instance, the path of a stock price is stochastic in nature, and one cannot reliably predict the exact stock price path with any certainty. However, the price evolution over time is enveloped in a process that generates these prices. *The process is fixed and predetermined, but the outcomes are not.* Hence, by stochastic simulation, we create multiple pathways of prices, obtain a statistical sampling of these simulations, and make inferences on the potential pathways that the actual price may undertake given the nature and parameters of the stochastic process used to generate the time series. Four stochastic processes are included in Risk Simulator's Forecasting tool, including Geometric Brownian Motion or random walk, which is the most common and prevalently used process due to its simplicity and wide-ranging applications. The other three stochastic processes are the mean-reversion process, jump-diffusion process, and a mixed process.

The interesting thing about stochastic process simulation is that historical data is not necessarily required; that is, the model does not have to fit any sets of historical data. Simply compute the expected returns and the volatility of the historical data or estimate them using comparable external data or make assumptions about these values.

Procedure

Run the stochastic forecast by using these two steps:

1. Start the module by selecting *Simulation* | *Forecasting* | *Stochastic Processes*.
2. Select the desired process, enter the required inputs, click on update chart a few times to make sure the process is behaving the way you expect it to, and click OK (Figure 8.9).

Results Interpretation

Figure 8.10 shows the results of a sample stochastic process. The chart shows a sample set of the iterations while the report explains the basics of stochastic processes. In addition, the forecast values (mean and standard

Regression Analysis Report

Regression Statistics

R-Squared (Coefficient of Determination)	0.6641
Adjusted R-Squared	0.4627
Multiple R (Multiple Correlation Coefficient)	0.7444
Standard Error of the Estimates (SEy)	14.7160
nObservations	23

The R-Squared or Coefficient of Determination indicates that of the variation in the dependent variable can be explained and accounted for by the independent variables in this regression analysis. However, in a multiple regression, the Adjusted R-Squared takes into account the existence of additional independent variables or regressors and adjusts this R-Squared value to a more accurate view of the regression's explanatory power. Hence, only of the variation in the dependent variable can be explained by the regressors.

The Multiple Correlation Coefficient (Multiple R) measures the correlation between the actual dependent variable (Y) and the estimated or fitted (Y) based on the regression equation. This is also the square root of the Coefficient of Determination (R-Squared.)

The Standard Error of the Estimates (SEy) describes the dispersion of data points above and below the regression line or plane. This value is used as part of the calculation to obtain the confidence interval of the estimates later.

Regression Results

	Intercept	X1	X2	X3	X4	X5
Coefficients	8.0795	–0.2440	0.1350	–0.0095	0.0175	0.4952
Standard Error	8.1593	0.4730	0.3051	0.1122	0.0418	1.5257
t-Statistic	0.9902	–0.6171	0.5415	–0.0845	0.4250	0.3044
p-Value	0.3325	0.6103	0.5934	0.9332	0.5743	0.7637
Lower 5%	–8.8419	–1.2268	–0.4890	–0.2422	–0.0888	–2.8783
Upper 95%	25.0009	0.7385	0.8007	0.2232	0.1044	3.8887

Degrees of Freedom	
Degrees of Freedom for Regression	6
Degrees of Freedom for Residual	22
Total Degrees of Freedom	27

Hypothesis Test	
Critical t-Statistic (99% confidence with df of 22)	2.8188
Critical t-Statistic (95% confidence with df of 22)	2.0730
Critical t-Statistic (90% confidence with df of 22)	1.7171

The Coefficients provide the estimated regression intercept and slopes. For instance, the coefficients are estimates of the true population – values in the following regression equation: $Y = 1_0 + -_1X_1 + 1_2X_2 + \_ + -_2X_2$. The Standard Error measures how accurate the predicted Coefficients are, and the t-Statistics are the ratios of each predicted Coefficient to its Standard Error.

The t-Statistic is used in hypothesis testing, where we set the null hypothesis (H_0) such that the real mean of the Coefficient = 0, and the alternate hypothesis (H_a) such that the real mean of the Coefficient is not equal to 0. A t-test is performed and the calculated t-Statistic is compared to the critical values at the relevant Degrees of Freedom for Residual. The t-test is very important as it calculates if each of the coefficients is statistically significant in the presence of the other regressors. This means that the t-test statistically verifies whether a regressor or independent variable should remain in the regression or it should be dropped.

The Coefficient is statistically significant if its calculated t-Statistic exceeds the Critical t-Statistic at the relevant degrees of freedom (df). The three main confidence levels used to test for significance are 90%, 95%, and 99%. If a Coefficient's t-Statistic exceeds the Critical level, it is considered statistically significant. Alternatively, the p-Value calculates each t-Statistic's probability of occurrence, which means that the smaller the p-Value, the more significant the Coefficient. The usual significant levels for the p-value are 0.01, 0.05, and 0.10, corresponding to the 99%, 95%, and 99% confidence levels.

The Coefficients with their p-values highlighted in blue indicate that they are statistically significant at the 95% confidence or 0.05 alpha level, while those highlighted in red indicate that they are not statistically significant at any of the alpha levels.

FIGURE 8.8 Multivariate regression results.

deviation) for each time period are provided. Using these values, you can decide which time period is relevant to your analysis, and set assumptions based on these mean and standard deviation values using the normal distribution. These assumptions can then be simulated in your own custom model.

Notes

Brownian Motion Random Walk Process The Brownian motion random walk process takes the form of

Analysis of Variance

	Sums of Squares	Mean of Squares	F-Statistic	P-Value
Regression	5919.2453	1153.0491	6.4073	0.0020
Residual	4783.7159	218.5327		
Total	10882.9543			

Hypothesis Test	
Critical t-Statistic (99% confidence with df of 4 and 3)	3.9880
Critical t-Statistic (95% confidence with df of 4 and 3)	2.5813
Critical t-Statistic (90% confidence with df of 4 and 3)	2.1279

The Analysis of Variance (ANOVA) table provides an F-test of the regression models overall statistical significance. Instead of looking at individual regressors as in the t-test, the F-test looks at all the estimated Coefficients statistical properties. The F-statistic is calculated as the ratio of the Regression's Mean of Squares to the Residual's Mean of Squares. The numerator measures how much of the regression is explained, while the denominator measures how much is unexplained. hence, the larger the F-statistic, the more significant the model. The corresponding p-Value is calculated to test the null hypothesis (H_0) where all the Coefficients are simultaneously equal to zero, versus the alternate hypothesis (H_a) that they are all simultaneously different from zero, indicating a significant overall regression model. If the p-Value is smaller than the 0.01, 0.05, or 0.10 alpha significance, then the regression is significant. The same approach can be applied to the F-statistic by comparing the calculated F-statistic with the critical F-values at various significance levels.

Forecasting

Period	Actual (Y)	Forecast (F)	Error (E)
1	10	16.7176	(6.7176)
2	13	18.1262	(6.1252)
3	14	19.9657	(5.9657)
4	15	22.1958	(7.1958)
5	18	23.6613	(5.6613)
6	6	24.8487	(18.84867)
7	87	24.7268	62.2732
8	21	24.9410	(3.9410)
9	23	25.9599	(2.9599)
10	34	25.8248	8.1752
11	26	27.1239	(1.1239)
12	28	27.9043	0.0957
13	29	31.0906	(2.0906)
14	30	34.3457	(4.3457)
15	33	28.9797	4.0203
16	23	36.2009	(13.2009)
17	39	37.2167	1.7833
18	44	46.1075	(2.1075)
19	44	43.8360	0.1640
20	46	48.3004	(2.3004)
21	48	48.3328	(0.3328)
22	55	53.6713	1.3287
23	57	54.3234	2.6766
24	66	67.1361	(1.1361)
25	48	48.3328	(0.3328)
26	55	53.6713	1.3287
27	57	54.3234	2.6766
28	66	67.1361	(1.1361)

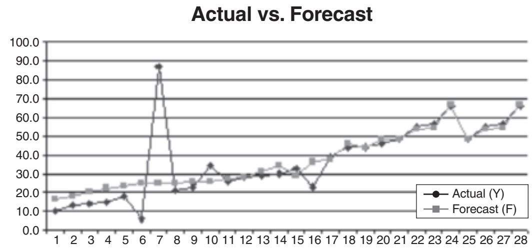


FIGURE 8.8 (Continued)

$$\frac{\delta S}{S} = \mu(\delta t) + \sigma\varepsilon\sqrt{\delta t}$$

for regular options simulation, or a more generic version takes the form of

$$\frac{\delta S}{S} = (\mu - \sigma^2\ /\ 2)\delta t + \sigma\varepsilon\sqrt{\delta t}$$

for a geometric process. For an exponential version, we simply take the exponentials, and as an example, we have

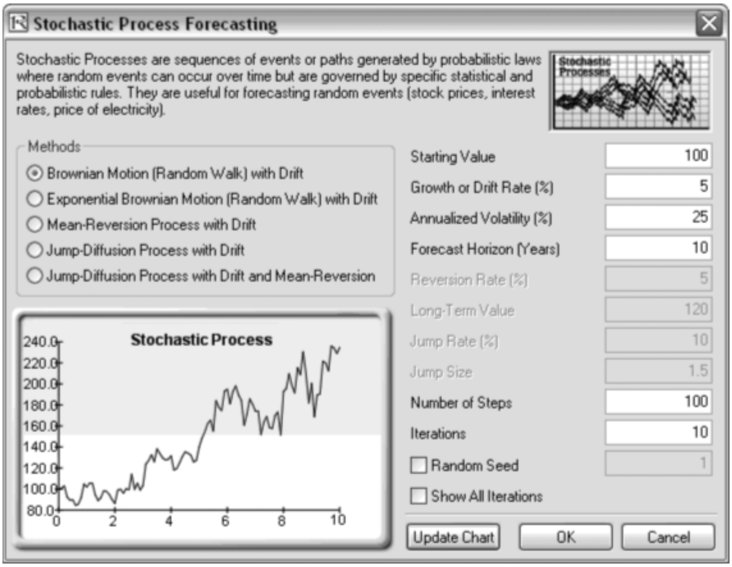


FIGURE 8.9 Stochastic process forecasting.

$$\frac{\delta S}{S} = \exp\left[\mu(\delta t) + \sigma\varepsilon\sqrt{\delta t}\right]$$

where S = the variable's previous value
δS = the change in the variable's value from one step to the next
μ = the annualized growth or drift rate
σ = the annualized volatility

To estimate the parameters from a set of time-series data, the drift rate and volatility can be found by setting μ to be the average of the natural logarithm of the relative returns $\ln\left[\frac{S_t}{S_{t-1}}\right]$ while σ is the standard deviation of all $\ln\left[\frac{S_t}{S_{t-1}}\right]$ values.

Mean-Reversion Process The following describes the mathematical structure of a mean-reverting process with drift:

$$\frac{\delta S}{S} = \eta(\bar{S}e^{\mu(\delta t)} - S)\delta t + \mu(\delta t) + \sigma\varepsilon\sqrt{\delta t}$$

Stochastic Process Forecasting

Statistical Summary

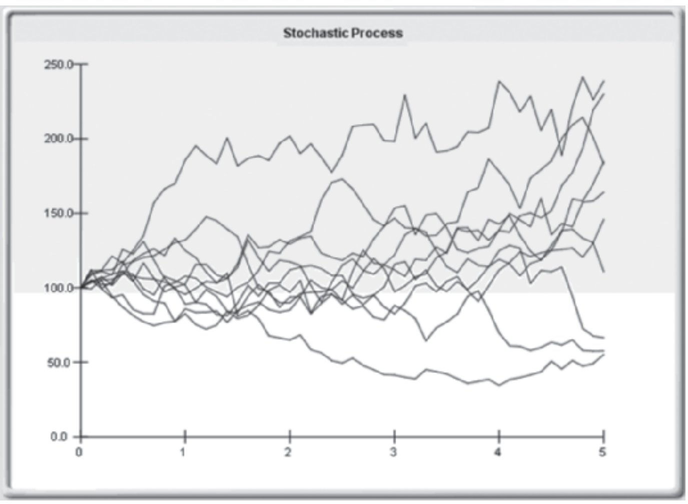
A stochastic process is a sequence of events or paths generated by probabilistic laws. That is, random events can occur over time but are governed by specific statistical and probabilistic rules. The main stochastic processes include Random Walk or Brownian Motion, Mean-Reversion, and Jump-Diffusion. These processes can be used to forecast a multitude of variables that seemingly follow random trends but yet are restricted by probabilistic laws.

The Random Walk Brownian Motion process can be used to forecast stock prices, prices of commodities, and other stochastic time-series data given a drift or growth rate and a volatility around the drift path. The Mean-Reversion process can be used to reduce the fluctuations of the Random Walk process by allowing the path to target a long-term value, making it useful for forecasting time-series variables that have a long-term rate such as interest rates and inflation rates (these are long-term target rates by regulatory authorities or the market). The Jump-Diffusion process is useful for forecasting time-series data when the variable can occasionally exhibit random jumps, such as oil prices or price of electricity (discrete exogenous event shocks can make prices jump up or down). Finally, these three stochastic processes can be mixed and matched as required.

The results on the right indicate the mean and standard deviation of all the iterations generated at each time step. If the Show All Iterations option is selected, each iteration pathway will be shown in a separate worksheet. the graph generated below shows a sample set of iteration pathways.

Stochastic Process: Brownian Motion (Random Walk) with Drift

Start Value	100	Steps	50.00	Jump Rate	N/A
Drift Rate	5.00%	Iterations	10.00	Jump Size	N/A
Volatility	25.00%	Reversion Rate	N/A	Random Seed	1720050445
Horizon	5	Long-Term Value	N/A		



Time	Mean	Stdev
0.0000	100.00	0.00
0.1000	106.32	4.05
0.2000	105.92	4.70
0.3000	105.23	8.23
0.4000	109.84	11.18
0.5000	107.57	14.67
0.6000	108.63	19.79
0.7000	107.85	24.18
0.8000	109.61	24.46
0.9000	109.57	27.99
1.0000	110.74	30.81
1.1000	111.53	35.05
1.2000	111.07	34.10
1.3000	107.52	32.85
1.4000	108.26	37.38
1.5000	106.36	32.19
1.6000	112.42	32.16
1.7000	110.08	31.24
1.8000	109.64	31.87
1.9000	110.18	36.43
2.0000	112.23	37.63
2.1000	114.32	33.10
2.2000	111.14	38.42
2.3000	111.03	37.69
2.4000	112.04	37.23
2.5000	112.98	40.84
2.6000	115.74	43.69
2.7000	115.11	43.64
2.8000	114.87	43.70
2.9000	113.28	42.25
3.0000	115.72	43.43
3.1000	120.05	50.48
3.2000	116.69	42.61
3.3000	118.31	45.57
3.4000	116.35	40.82
3.5000	115.71	40.33
3.6000	118.69	41.45
3.7000	121.66	45.34
3.8000	121.40	45.03
3.9000	125.19	48.19
4.0000	129.65	55.44
4.1000	129.61	53.82
4.2000	125.86	49.68
4.3000	125.70	53.79
4.4000	126.72	49.70
4.5000	129.52	50.28
4.6000	132.28	49.70
4.7000	138.47	56.77
4.8000	139.69	66.32
4.9000	140.85	65.95
5.0000	143.61	68.65

FIGURE 8.10 Stochastic forecast result.

In order to obtain the rate of reversion and long-term rate, using the historical data points, run a regression such that $Y_t - Y_{t-1} = \beta_0 + \beta_1 Y_{t-1} + \varepsilon$ and we find $\eta = -\ln[1 + \beta_1]$ and $\bar{S} = -\beta_0/\beta_1$, where

η = the rate of reversion to the mean
$\bar{S}$ = the long-term value the process reverts to

Y = the historical data series
β_0 = the intercept coefficient in a regression analysis
β_1 = the slope coefficient in a regression analysis

Jump Diffusion Process A jump diffusion process is similar to a random walk process, but there is a probability of a jump at any point in time. The occurrences of such jumps are completely random, but the probability and magnitude are governed by the process itself.

$$\frac{\delta S}{S} = \eta(\bar{S}e^{\mu(\delta t)} - S)\delta t + \mu(\delta t) + \sigma\varepsilon\sqrt{\delta t} + \theta F(\lambda)(\delta t)$$

for a jump diffusion process, where

θ = the jump size of S
$F(\lambda)$ = the inverse of the Poisson cumulative probability distribution
λ = the jump rate of S

The jump size can be found by computing the ratio of the postjump to the prejump levels, and the jump rate can be imputed from past historical data. The other parameters are found the same way as shown previously.

NONLINEAR EXTRAPOLATION

Theory

Extrapolation involves making statistical forecasts by using historical trends that are projected for a specified period of time into the future. It is only used for time-series forecasts. For cross-sectional or mixed panel data (time series with cross-sectional data), multivariate regression is more appropriate. This methodology is useful when major changes are not expected; that is, causal factors are expected to remain constant or when the causal factors of a situation are not clearly understood. It also helps discourage the introduction of personal biases into the process. Extrapolation is fairly reliable, relatively simple, and inexpensive. However, extrapolation, which assumes that recent and historical trends will continue, produces large forecast errors if discontinuities occur within the projected time period; that is, pure extrapolation of time series assumes that all we need to know is contained in the historical values of the series being forecasted. If we assume that past behavior is a good predictor of future behavior, extrapolation is appealing. This makes it a useful approach when all that is needed are many short-term forecasts.

This methodology estimates the *f(x)* function for any arbitrary x value, by interpolating a smooth nonlinear curve through all the x values, and using this smooth curve, extrapolates future x values beyond the historical data set. The methodology employs either the polynomial functional form or the rational functional form (a ratio of two polynomials). Typically, a polynomial functional form is sufficient for well-behaved data; however, rational functional forms are sometimes more accurate (especially with polar functions, i.e., functions with denominators approaching zero).

Procedure

Use the following steps to run a nonlinear extrapolation:

1. Start Excel and enter your data or open an existing worksheet with historical data to forecast (the illustration shown next uses the file *Nonlinear Extrapolation* from the examples folder).
2. Select the time-series data and select *Simulation* | *Forecasting* | *Nonlinear Extrapolation*.
3. Select the extrapolation type (automatic selection, polynomial function, or rational function are available, but in this example, use automatic selection) and enter the number of forecast period desired (Figure 8.11), and click OK.

Historical Sales Revenues
Polynomial Growth Rates

Year	Month	Period	Sales
2004	1	1	$1.00
2004	2	2	$6.73
2004	3	3	$20.52
2004	4	4	$45.25
2004	5	5	$83.59
2004	6	6	$138.01
2004	7	7	$210.87
2004	8	8	$304.44
2004	9	9	$420.89
2004	10	10	$562.34
2004	11	11	$730.85
2004	12	12	$928.43

Historical Net Income
Sinusoidal Growth Rates

Year	Month	Period	Income
2004	1	1	$84.15
2004	2	2	$90.93
2004	3	3	$14.11
2004	4	4	($75.68)
2004	5	5	($95.89)
2004	6	6	($27.94)
2004	7	7	$65.70
2004	8	8	$98.94
2004	9	9	$41.21
2004	10	10	($54.40)
2004	11	11	($100.00)
2004	12	12	($53.66)
2005	1	13	$42.02
2005	2	14	$99.06
2005	3	15	$65.03
2005	4	16	($28.79)
2005	5	17	($96.14)
2005	6	18	($75.10)

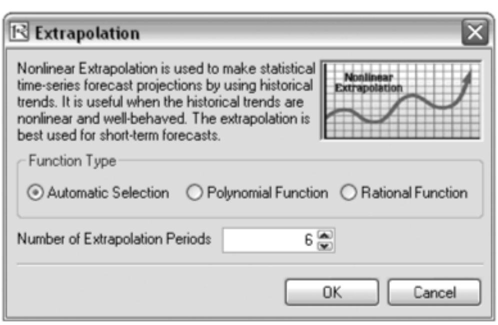


FIGURE 8.11 Running a nonlinear extrapolation.

Results Interpretation

The results report shown in Figure 8.12 shows the extrapolated forecast values, the error measurements, and the graphical representation of the extrapolation results. The error measurements should be used to check the validity of the forecast and are especially important when used to compare the forecast quality and accuracy of extrapolation versus time-series analysis.

Notes

When the historical data is smooth and follows some nonlinear patterns and curves, extrapolation is better than time-series analysis. However, when the data patterns follow seasonal cycles and a trend, time-series analysis will

Nonlinear Extrapolation

Statistical Summary

Extrapolation involves making statistical projections by using historical trends that are projected for a specified period of time into the future. It is only used for time-series forecasts. for cross-sectional or mixed panel data (time-series with cross-sectional data), multivariate regression is more appropriate. This methodology is useful when major changes are not expected, that is, causal factors are expected to remain constant or when the causal factors of a situation are not clearly understood. It also helps discourage introduction of personal biases into the process. Extrapolation is fairly reliable, relatively simple, and inexpensive. However, extrapolation, which assumes that recent and historical trends will continue, produces large forecast errors if discontinuities occur within the projected time period. That is, pure extrapolation of time series assumes that all we need to know is contained in the historical values of the series being forecasted. If we assume that past behavior is a good predictor of future behavior, extrapolation is appealing. This makes it a useful approach when all that is needed are many short-term forecasts.

This methodology estimates the *f(x)* function for any arbitrary *x* value, by interpolating a smooth nonlinear curve through all the *x* values, and using this smooth curve, extrapolates future *x* values beyond the historical data set. the methodology employs either the polynomial functional form or the rational functional form (a ratio of two polynomials). Typically, a polynomial functional form is sufficient for well-behaved data; however, rational functional forms are sometimes more accurate (especially with polar functions, i.e., functions with denominators approaching zero).

Period	Actual	Forecast Fit	Estimate Error
1	1.00		
2	6.73	1.00	
3	20.52	–1.42	–8.15
4	45.25	99.82	119.36
5	83.59	55.92	–46.67
6	138.01	136.71	14.39
7	210.87	211.96	1.69
8	304.44	304.43	–0.41
9	420.89	420.89	0.01
10	562.34	562.34	0.00
11	730.85	730.85	0.00
12	928.43	928.43	0.00
Forecast 13		1157.03	0.00
Forecast 14		1418.57	0.00
Forecast 15		1714.95	0.00
Forecast 16		2048.00	0.00
Forecast 17		2419.55	0.00
Forecast 18		2831.39	0.00

Error Measurements	
RMSE	19.6799
MSE	387.2974
MAD	10.2095
MAPE	31.56%
Theil's U	1.1210

Function Type: Rational

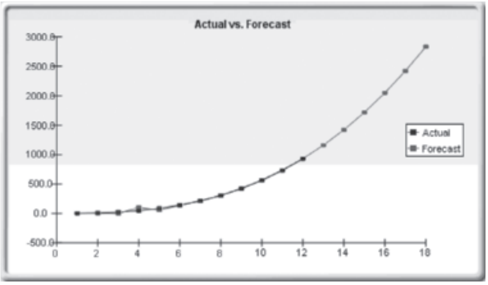


FIGURE 8.12 Nonlinear extrapolation results.

provide better results. It is always advisable to run both time-series analysis and extrapolation and compare the results to see which has a lower error measure and a better fit.

BOX–JENKINS ARIMA ADVANCED TIME-SERIES

Theory

One very powerful advanced times-series forecasting tool is the ARIMA or autoregressive integrated moving average approach, which assembles three separate tools into a comprehensive model. The first tool segment is the autoregressive or "AR" term, which corresponds to the number of lagged value of the residual in the unconditional forecast model. In essence, the model captures the historical variation of actual data to a forecasting model and uses this variation or residual to create a better predicting model. The second tool segment is the integration order or the "I" term. This integration term corresponds to the number of differencing the time-series to be forecasted goes through to make the time-series data stationary. This element accounts for any nonlinear growth rates existing in the data. The third tool segment is the moving average or "MA" term, which is essentially the moving average of lagged forecast errors. By incorporating these lagged forecast errors, the model in essence learns from its forecast errors or mistakes and corrects for them through a moving average calculation. The ARIMA model follows the Box–Jenkins methodology with each term representing steps taken in the model construction until only random noise remains. Also, ARIMA modeling uses correlation techniques in generating forecasts. ARIMA can be used to model patterns that may not be visible in plotted data. In addition, ARIMA models can be mixed with exogenous variables, but make sure that the exogenous variables have enough data points to cover the additional number of periods to forecast.

There are many reasons why an ARIMA model is superior to common time-series analysis and multivariate regressions. The common finding in time-series analysis and multivariate regression is that the error residuals are correlated with their own lagged values. This serial correlation violates the standard assumption of regression theory that disturbances are not correlated with other disturbances. The primary problems associated with serial correlation are:

- Regression analysis and basic time-series analysis are no longer efficient among the different linear estimators. However, as the error residuals can help to predict current error residuals, we can take advantage of this information to form a better prediction of the dependent variable using ARIMA.

- Standard errors computed using the regression and time-series formula are not correct and are generally understated. If there are lagged dependent variables set as the regressors, regression estimates are biased and inconsistent but can be fixed using ARIMA.

Autoregressive integrated moving average or ARIMA(p,d,q) models are the extension of the AR model that uses three components for modeling the serial correlation in the time-series data. The first component is the autoregressive (AR) term. The AR(p) model uses the p lags of the time series in the equation. An AR(p) model has the form: $y_t = a_1 y_{t-1} + \ldots + a_p y_{t-p} + e_t$. The second component is the integration (d) order term. Each integration order corresponds to differencing the time series. I(1) means differencing the data once. I(d) means differencing the data d times. The third component is the moving average (MA) term. The MA(q) model uses the q lags of the forecast errors to improve the forecast. An MA(q) model has the form: $y_t = e_t + b_1 e_{t-1} + \ldots + b_q e_{t-q}$. Finally, an ARMA(p,q) model has the combined form: $y_t = a_1 y_{t-1} + \ldots + a_p y_{t-p} + e_t + b_1 e_{t-1} + \ldots + b_q e_{t-q}$.

Procedure

To run a Box–Jenkins ARIMA model, follow these steps:

1. Start Excel and enter your data or open an existing worksheet with historical data to forecast (the illustration shown next uses the example file *Time-Series Forecasting*).
2. Select *Simulation* | *Forecasting* | *ARIMA* and select the time-series data.
3. Enter the relevant *p*, *d*, and *q* parameters (positive integers only), enter the number of forecast periods desired, and click OK.

Results Interpretation

In interpreting the results of an ARIMA model, most of the specifications are identical to the multivariate regression analysis (see Chapter 9, Using the Past to Predict the Future, for more technical details about interpreting the multivariate regression analysis and ARIMA models). However, several additional sets of results are specific to the ARIMA analysis as seen in Figure 8.13. The first is the addition of Akaike Information Criterion (AIC) and Schwarz Criterion (SC), which are often used in ARIMA model selection and identification. That is, AIC and SC are used to determine if a particular model with a specific set of *p*, *d*, and *q* parameters is a good statistical fit. SC imposes a greater penalty for additional coefficients than the AIC, but generally, the model with the lowest AIC and SC values should be chosen. Finally, an additional set of results called the autocorrelation (AC) and partial autocorrelation (PAC) statistics are provided in the ARIMA report.

ARIMA (Autoregressive Integrated Moving Average)

Regression Statistics

R-Squared (Coefficient of Determination)	0.7708	Akaike Information Criterion (AIC)	14.2506
Adjusted R-Squared	0.7573	Schwarz Criterion (SC)	14.3500
Multiple R (Multiple Correlation Coefficient)	0.8779	Log Likelihood	–133.3807
Standard Error of the Estimates (SEy)	580.9368	Durbin-Watson Statistic	2.3576
nObservations	19	Number of Iterations	0

Autoregressive integrated moving average (ARIMA (p, d, q)) models are the extension of the AR model that use three components fro modeling the serial correlation in the time series data. The first component is the autoregressive (AR) term. The AR (p) model uses the p lags of the time series in the equation. An AR (p) model has the form: $y(t)=a(1)*y(t-1)+...+a(p)*y(t-p)+e(t)$. The second component is the integration (d) order term. Each integration order corresponds to differencing the time series. I(1) means differencing the data once. I(d) means differencing the data d times. The third component is the moving average (MA) term. The MA (q) model uses the q lags of the forecast errors to improve the forecast. An MA (q) model has the form: $y(t)=e(t)+b(1)*e(t-1)+...+b(q)*e(t-q)$. Finally, an ARMA (p, q) model has the combined form: $y(t)=a(1)*y(t-1)+...+a(p)*y(t-p)+e(t)+b(1)*e(t-1)+...+b(q)*e(t-q)$.

The R-Squared or Coefficient of Determination indicates that of the variation in the dependent variable can be explained and accounted for by the independent variables in this regression analysis. However, in a multiple regression, the Adjusted R-Squared takes into account the existence of additional independent variables of regressors and adjusts this R-Squared value to a more accurate view of the regressions's explanatory power. Hence, only of the variation in the dependent variable can be explained by the regressors. However, under some circumstances, it tends to be unreliable.

The Multiple Correlation Coefficient (Multiple R) measures the correlation between the actual dependent variable (*Y*) and the estimated or fitted (Y) based on the regression equation. This is also the square root of the Coefficient of Determination (R-Squared).

The Standard Error of the Estimates (SEy) describes the dispersion of data points above and below the regression line or plane. This value is used as part of the calculation to obtain the confidence interval of the estimates later.

The AIC and SC are often used in model selection. SC imposes a greater penalty for additional coefficient. Generally, the user should select a model with the lowest value of the AIC and SC.

The Durbin-Watson statistic measures the serial correlation in the residuals. Generally, DW less than 2 implies positive serial correlation.

Regression Results

	Intercept	Y(–1)
Coefficients	116.3328	0.9895
Standard Error	179.9049	0.1309
t-Statistic	0.6466	7.5604
p-Value	0.5265	0.0000
Lower 5%	–263.2333	0.7134
Upper 95%	495.8989	1.2656

Degrees of Freedom	
Degrees of Freedom for Regression	1
Degrees of Freedom for Residual	17
Total Degrees of Freedom	18

Hypothesis Test	
Critical t-Statistic (99% confidence with df of 22)	63.6567
Critical t-Statistic (95% confidence with df of 22)	2.1098
Critical t-Statistic (90% confidence with df of 22)	1.7341

The Coefficients provide the estimated regression intercept and slopes. For instance, the coefficients are the b values in the following regression equation: $Y = b(0) + b(1)\times(1) + b(2)\times(2) + ... + b(n)\times(n)$. The Standard Errors measure how accurate the predicted Coefficients are, and the t-Statistics are the ratios of each predicted Coefficient to its Standard Error.

The t-Statistic is used in hypothesis testing, where we set the null hypothesis (H_0) such that the real mean of the Coefficient = 0, and the alternate hypothesis (H_a) such that the real mean of the Coefficient is not equal to 0. A t-test is performed and the calculated t-Statistic is compared to the critical values at the relevant Degrees of Freedom for Residual. The t-test is very important as it calculates if each of the coefficients is statistically significant in the presence of the other regressors. This means that the t-test statistically verifies whether a regressor or independent variable should remain in the regression or it should be dropped.

The Coefficient is statistically significant if its calculated t-Statistic exceeds the Critical t-Statistic at the relevant degrees of freedom (df). The three main confidence levels used to test for significance are 90%, 95%, and 99%. If a Coefficient's t-Statistic exceeds the Critical level, it is considered statistically significant. Alternatively, the p-Value calculates each t-Statistic's probability of occurrence, which means that the smaller the p-value, the more significant the Coefficient. The usual critical levels for the p-value are 0.01, 0.05, and 0.10, corresponding to the 99%, 95%, and 99% confidence levels.

The Coefficients with their p-values highlighted in blue indicate that they are statistically significant at the 95% confidence or 0.05 alpha level, while those highlighted in red indicate that they are not statistically significant at any of the alpha levels.

(*continued*)

FIGURE 8.13 Box–Jenkins ARIMA forecast report.

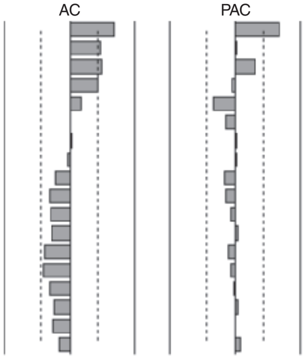
Analysis of Variance

	Sums of Squares	Mean of Squares	F-Statistic	P-Value	Hypothesis Test	
Regression	4682236.0689	4682238.0689	57.1604	0.0000	Critical t-Statistic (99% confidence with df of 1 and 17)	8.3997
Residual	1392538.5521	81914.0325			Critical t-Statistic (95% confidence with df of 1 and 17)	4.4513
Total	6074776.6211	4764152.1014			Critical t-Statistic (90% confidence with df of 1 and 17)	3.0262

The Analysis of Variance (ANOVA) table provides an F-test of the regression model's overall statistical significance. Instead of looking at individual regressors as in the t-test, the F-test looks at all the estimated Coefficients statistical properties. The F-statistic is calculated as the ratio of the Regression's Mean of Squares to the Residual's Mean of Squares. The numerator measures how much of the regression is explained, while the denominator measures how much is unexplained. Hence, the larger the F-statistic, the more significant the model. The corresponding p-value is calculated to test the null hypothesis (H_0) where all the Coefficients are simultaneously equal to zero, versus the alternate hypothesis (H_a) that they are all simultaneously different from zero, indicating a significant overall regression model. If the p-value is smaller than the 0.01, 0.05, or 0.10 alpha significance, then the regression is significant. The same approach can be applied to the F-statistic.

Autocorrelation

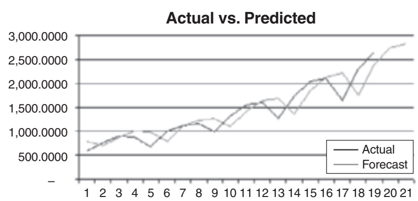
Time Log	AC	PAC	LBound	UBound	Q-Stat	Prob
1	0.6871	0.6871	(0.4472)	0.4472	10.4657	0.0012
2	0.4850	0.0244	(0.4472)	0.4472	15.9865	0.0003
3	0.5045	0.3083	(0.4472)	0.4472	22.3339	0.0001
4	0.4334	(0.0512)	(0.4472)	0.4472	27.3303	0.0000
5	0.1720	(0.3282)	(0.4472)	0.4472	28.1730	0.0000
6	0.0185	(0.1400)	(0.4472)	0.4472	28.1835	0.0001
7	0.0243	0.0334	(0.4472)	0.4472	28.2032	0.0002
8	(0.0280)	0.0286	(0.4472)	0.4472	28.2316	0.0004
9	(0.2099)	(0.1544)	(0.4472)	0.4472	29.9697	0.0004
10	(0.3074)	(0.1478)	(0.4472)	0.4472	34.1800	0.0002
11	(0.2828)	(0.0666)	(0.4472)	0.4472	38.1679	0.0001
12	(0.2734)	0.0529	(0.4472)	0.4472	42.4282	0.0000
13	(0.3774)	(0.0941)	(0.4472)	0.4472	51.9000	0.0000
14	(0.4018)	(0.0644)	(0.4472)	0.4472	64.7818	0.0000
15	(0.2998)	(0.0012)	(0.4472)	0.4472	73.7471	0.0000
16	(0.2303)	0.0428	(0.4472)	0.4472	80.8003	0.0000
17	(0.2489)	0.0064	(0.4472)	0.4472	93.1562	0.0000
18	(0.1652)	0.0892	(0.4472)	0.4472	104.0461	0.0000



If autocorrelation AC(1) is nonzero, it means that the series is first order serially correlated. If AC(k) dies off more or less geometrically with increasing lag, it implies that the series follows a low-order autoregressive process. If AC(k) drops to zero after a small number of lags, it implies that the series follows a low-order moving average process. Partial correlation (PAC(k) measures the correlation of values that are k periods apart after removing the correlation from the intervening lags. If the pattern of autocorrelation can be captured by an autoregression of order less than k, then the partial autocorrelation at lag k will be close to zero. Ljung-Box Q-statistics and their p-values at lag k has the null hypothesis that there is no autocorrelation up to order k. The dotted lines in the plots of the autocorrelations are the approximate two standard error bounds. If the autocorrelation is within these bounds, it is not significantly different from zero at (approximately) the 5% significance level.

Forecasting

Period	Actual (Y)	Forecast (F)	Error (E)
1	584.1000	793.3540	(209.2540)
2	765.4000	694.3043	71.0957
3	892.3000	873.7021	18.5979
4	885.4000	999.2706	(113.8706)
5	677.0000	992.4430	(315.4430)
6	1,006.6000	786.2296	220.3704
7	1,122.1000	1,112.3713	9.7267
8	1,163.4000	1,226.6595	(63.2595)
9	993.2000	1,267.5262	(274.3262)
10	1,312.5000	1,099.1119	213.3881
11	1,545.3000	1,415.0618	130.2382
12	1,596.2000	1,645.4192	(49.2192)
13	1,260.4000	1,695.7852	(435.3852)
14	1,735.2000	1,363.5084	371.6916
15	2,029.7000	1,833.3267	196.3733
16	2,107.8000	2,124.7368	(16.9368)
17	1,650.3000	2,202.0173	(551.7173)
18	2,304.4000	1,749.3175	555.0825
19	2,639.4000	2,396.5546	242.8454
20		2,728.0397	
21		2,815.7494	



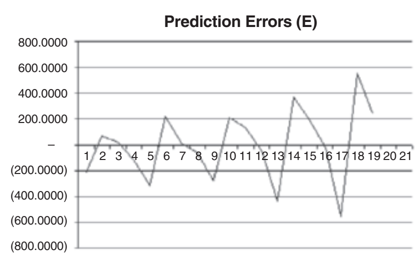


FIGURE 8.13 (Continued)

For instance, if autocorrelation AC(1) is nonzero, it means that the series is first order serially correlated. If AC dies off more or less geometrically with increasing lags, it implies that the series follows a low-order autoregressive process. If AC drops to zero after a small number of lags, it implies that the series follows a low-order moving-average process. In contrast, PAC measures the correlation of values that are *k* periods apart after removing the correlation from the intervening lags. If the pattern of autocorrelation can be captured by an autoregression of order less than *k*, then the partial autocorrelation at lag *k* will be close to zero. The Ljung–Box Q-statistics and their p-values at lag *k* are also provided, where the null hypothesis being tested is such that there is no autocorrelation up to order *k*. The dotted lines in the plots of the autocorrelations are the approximate two standard error bounds. If the autocorrelation is within these bounds, it is not significantly different from zero at approximately the 5 percent significance level. Finding the right ARIMA model takes practice and experience. These AC, PAC, SC, and AIC are highly useful diagnostic tools to help identify the correct model specification. Finally, the ARIMA parameter results are obtained using sophisticated optimization and iterative algorithms, which means that although the functional forms look like those of a multivariate regression, they are not the same. ARIMA is a much more computationally intensive and advanced econometric approach.

QUESTIONS

1. What are the differences between time-series forecasting techniques and nonlinear extrapolation?
2. Which forecasting method requires existing data and which method does not?
3. How do you use the software to perform qualitative forecasts?
4. Replicate all the examples in this chapter.
5. Time-series data that exhibit seasonality are easier to forecast than data that exhibit cyclicality. Is this statement true and why or why not?

The following pages present additional hands-on exercises on forecasting and review all the techniques covered in this chapter.

Time-Series Forecasting

This sample model is used to illustrate how to use Risk Simulator for:

1. Running a Time-Series Forecast

Model Background
File Name: Time-Series Forecasting.xls
The historical sales revenue data are located in the Time-Series Data worksheet. The data are quarterly sales revenue from Q1 2000 to Q4 2004. The data exhibits quarterly seasonality, which means that the seasonality is 4 (there are 4 quarters in 1 year or 1 cycle). The sample data set is seen below.

Time-Series Analysis
To run this model, simply:
1. Select the historical data (cells *H11:H30*)
2. Select **Simulation | Forecasting | Time-Series Analysis**
3. Select *Auto Model Selection* , *Forecast 4 Periods* and *Seasonality 4 Periods*

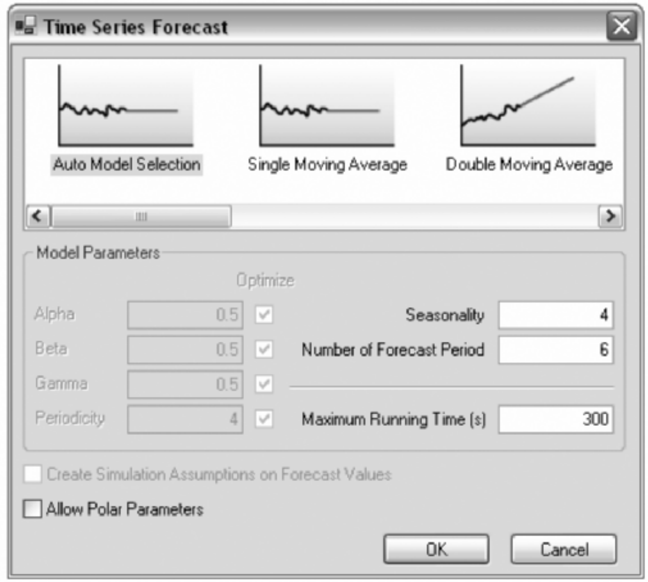
Note that you can only select Create Simulation Assumptions if an existing Simulation Profile exists (if not, click on Simulation, New Simulation Profile, and then run the time-series forecast per the steps above but remember to check the Create Simulation Assumptions box).

Model Results Analysis

For your convenience, the analysis Report and Methodology sheets are included. A fitted chart and forecast values are provided in the report as well as the error measures and a statistical summary of the methodology. The Methodology sheet provides the statistical results from all 8 time-series methodologies.

Historical Sales Revenues

Year	Quarter	Period	Sales
2000	1	1	$684.20
2000	2	2	$584.10
2000	3	3	$765.40
2000	4	4	$892.30
2001	1	5	$885.40
2001	2	6	$677.00
2001	3	7	$1,006.60
2001	4	8	$1,122.10
2002	1	9	$1,163.40
2002	2	10	$993.20
2002	3	11	$1,312.50
2002	4	12	$1,545.30
2003	1	13	$1,596.20
2003	2	14	$1,260.40
2003	3	15	$1,735.20
2003	4	16	$2,029.70
2004	1	17	$2,107.80
2004	2	18	$1,650.30
2004	3	19	$2,304.40
2004	4	20	$2,639.40



Holt-Winter's Multiplicative

Summary Statistics

Alpha, Beta, Gamma	RMSE	Alpha, Beta, Gamma	RMSE
0.00, 0.00, 0.00	914.824	0.00, 0.00, 0.00	914.824
0.10, 0.10, 0.10	415.322	0.10, 0.10, 0.10	415.322
0.20, 0.20, 0.20	187.202	0.20, 0.20, 0.20	187.202
0.30, 0.30, 0.30	118.795	0.30, 0.30, 0.30	118.795
0.40, 0.40, 0.40	101.794	0.40, 0.40, 0.40	101.794
0.50, 0.50, 0.50	102.143		

The analysis was run with alpha = 0.2429, beta = 1.0000, gamma = 0.7797, and seasonality = 4

Time-Series Analysis Summary

When both seasonality and trend exist, more advanced models are required to decompose the data into their base elements: a base-case level (L) weighted by the alpha parameter; a trend component (b) weighted by the beta parameter; and a seasonality component (S) weighted by the gamma parameter. Several methods exist but the two most common are the Holt-Winters' additive seasonality and Holt-Winters' multiplicative seasonality methods. In the Holt-Winter's additive model, the base case level, seasonality, and trend are added together to obtain the forecast fit.

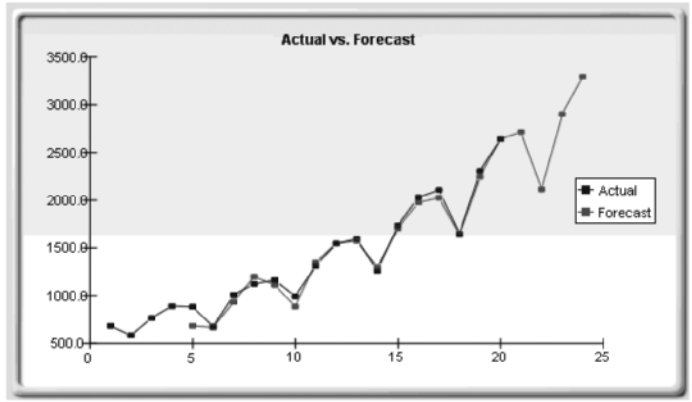
The best-fitting test for the moving average forecast uses the root mean squared errors (RMSE). The RMSE calculates the square root of the average squared deviations of the fitted values versus the actual data points.

Mean Squared Error (MSE) is an absolute error measure that squares the errors (the difference between the actual historical data and the forecast-fitted data predicted by the model) to keep the positive and negative errors from canceling each other out. This measure also tends to exaggerate large errors by weighting the large errors more heavily than smaller errors by squaring them, which can help when comparing different time-series models. Root Mean Square Error (RMSE) is the square root of MSE and is the most popular error measure, also known as the quadratic loss function. RMSE can be defined as the average of the absolute values of the forecast errors and is highly appropriate when the cost of the forecast errors is proportional to the absolute size of the forecast error. The RMSE is used as the selection criteria for the best-fitting time-series model.

Mean Absolute Percentage Error (MAPE) is a relative error statistic measured as an average percent error of the historical data points and is most appropriate when the cost of the forecast error is more closely related to the percentage error than the numerical size of the error. Finally, an associated measure is the Theil's U statistic, which measures the naivety of the model's forecast. That is, if the Theil's U statistic is less than 1.0, then the forecast method used provides an estimate that is statistically better than guessing.

Period	Actual	Forecast Fit
1	684.20	
2	584.10	
3	765.40	
4	892.30	
5	885.40	684.20
6	677.00	667.55
7	1006.60	935.45
8	1122.10	1198.09
9	1163.40	1112.48
10	993.20	887.95
11	1312.50	1348.38
12	1545.30	1546.53
13	1596.20	1572.44
14	1260.40	1299.20
15	1735.20	1704.77
16	2029.70	1976.23
17	2107.80	2026.01
18	1650.30	1637.28
19	2304.40	2245.93
20	2639.40	2643.09
Forecast 21		2713.69
Forecast 22		2114.79
Forecast 23		2900.42
Forecast 24		3293.81

Error Measurements	
RMSE	71.8132
MSE	5157.1348
MAD	53.4071
MAPE	4.50%
Theil's U	0.3054



Multiple Regression

This sample model is used to illustrate how to use Risk Simulator for:

1. Running a Multiple Regression Analysis

Model Background

File Name: Multiple Regression.xls

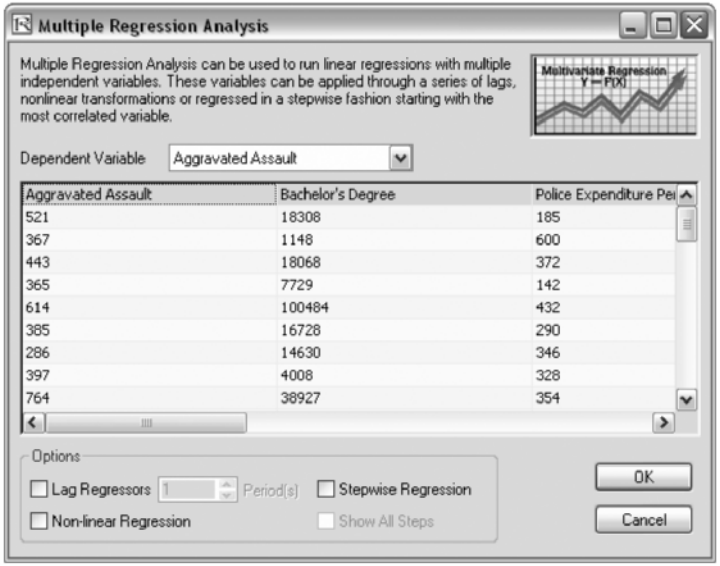
This example shows how a multiple regression can be run using Risk Simulator. The raw data are arranged in the Cross Sectional Data worksheet and contains cross-sectional data on all 50 U.S. states on the number of aggravated assaults (in thousands) per year, the number of bachelor's degrees awarded per year, police expenditure per capita population, population size in millions, population density (person per square mile), and unemployment rate. The idea is to see if there is a relationship between the number of aggravated assaults per year and these explanatory variables using multiple regression analysis. A sample snapshot of the data set is seen below:

Aggravated Assault	Bachelor's Degree	Police Expenditure Per Capita	Population in Millions	Population Density (Person/Sq Mile)	Unemployment Rate
521	18308	185	4.041	79.6	7.2
367	1148	600	0.55	1	8.5
443	18068	372	3.665	32.3	5.7
365	7729	142	2.351	45.1	7.3
614	100484	432	29.76	190.8	7.5
385	16728	290	3.294	31.8	5
286	14630	346	3.287	678.4	6.7
397	4008	328	0.666	340.8	6.2
764	38927	354	12.938	239.6	7.3

Multiple Regression Analysis

To run this model, simply:

1. In the *Cross Sectional Data worksheet* , select the area *C5:H55*.
2. Select **Simulation | Forecasting | Multiple Regression**.
3. Choose *Aggravated Assault* as the dependent variable in the regression and click on **OK**.



Results Summary

Refer to the Report worksheet for details on the regression output. The report has more details on the interpretation of specific statistical results. The report provides the following elements: multiple regression and analysis of variance output, including coefficients of determination, hypothesis test results (single variable **t**-test and multiple variable F-test), computed coefficients for each regressor, fitted chart, and much more.

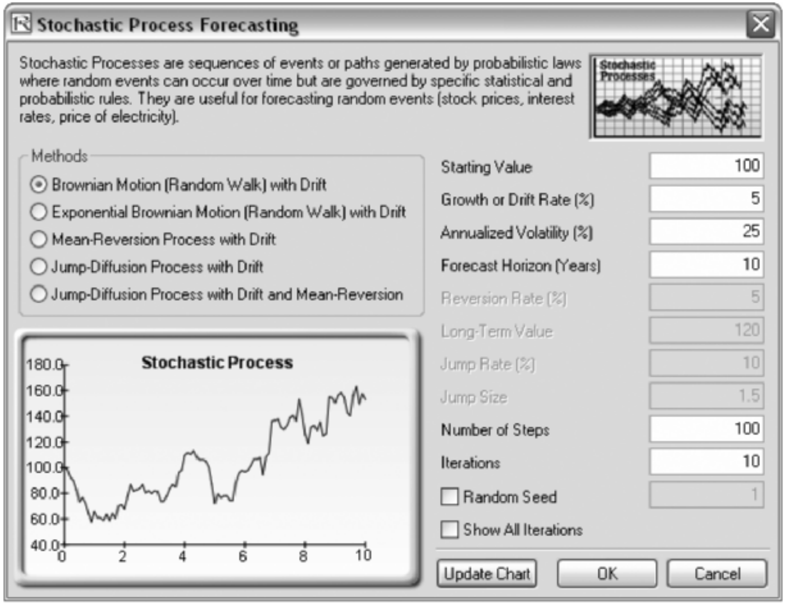
Stochastic Processes

This sample model is used to illustrate how to use Risk Simulator for:

1. Simulating Stochastic Processes: Brownian Motion Random Walk, Mean-Reversion, Jump-Diffusion, and Mixed Models.

Model Background
File Name: Stochastic Processes.xls
A stochastic process is a sequence of events or paths generated by probabilistic laws. That is, random events can occur over time but are governed by specific statistical and probabilistic rules. The main stochastic processes include Random Walk or Brownian Motion, Mean-Reversion and Jump-Diffusion. These processes can be used to forecast a multitude of variables that seemingly follow random trends but yet are restricted by probabilistic laws. We can use Risk Simulator's Stochastic Process module to simulate and create such processes. These processes can be used to forecast a multitude of time-series data including stock prices, interest rates, inflation rates, oil prices, eletricity prices, commodity prices, and so forth.

Stochastic Process Forecasting
To run this model, simply:
1. Select **Simulation | Forecasting | Stochastic Processes**
2. Enter a set of relevant inputs or use the existing inputs as a test case
3. Select the relevant process to simulate
4. Click on **Update Chart** to view the updated computation of a single path or click **OK** to create the process



Model Results Analysis

For your convenience, the analysis Report sheet is included. A stochastic time-series chart and forecast values are provided in the report. Each step's time period, mean and standard deviation of the forecast is provided in the report. The mean values can be used as the single point estimate or assumptions can be manually generated for the desired time period. That is, finding the appropriate time period, create an assumption with a Normal distribution with the appropriate mean and standard deviation computed. A sample chart with 10 iteration paths is included to graphically illustrate the behavior of the forecasted process.

Stochastic Process Forecasting

Statistical Summary

A stochastic process is a sequence of events or paths generated by probabilistic laws. That is, random events can occur over time but are governed by specific statistical and probabilistic rules. The main stochastic processes include Random Walk or Brownian Motion, Mean-Reversion, and Jump-Diffusion. These processes can be used to forecast a multitude of variables that seemingly follow random trends but yet are restricted by probabilistic laws.

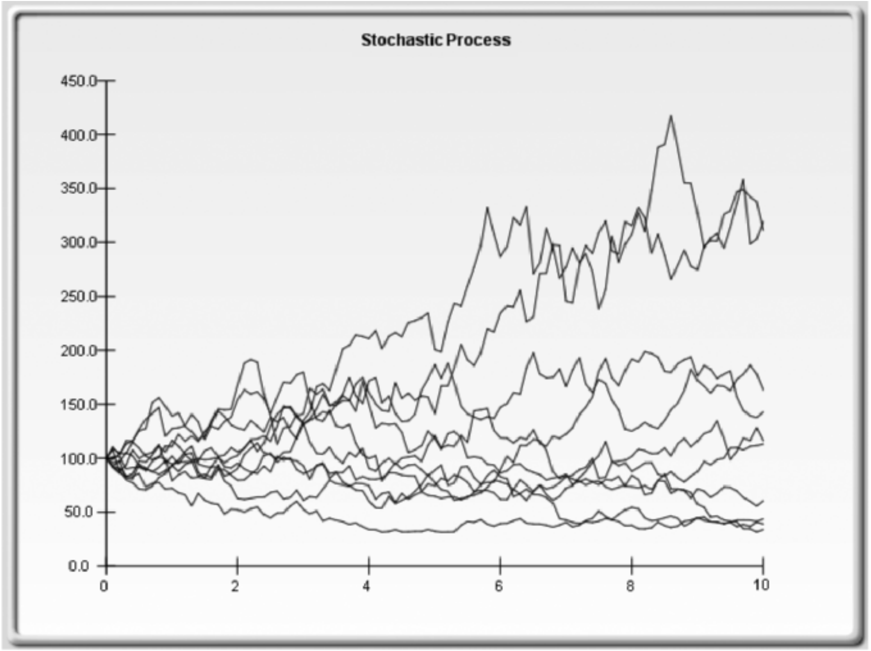
The Random Walk Brownian Motion process can be used to forecast stock prices, prices of commodities, and other stochastic time-series data given a drift or growth rate and a volatility around the drift path. The Mean-Reversion process can be used to reduce the fluctuations of the Random Walk process by allowing the path to target a long-term value, making it useful for forecasting time-series variables that have a long-term rate such as interest rates and inflation rates (these are long-term target rates by regulatory authorities or the market). The Jump-Diffusion process is useful for forecasting time-series data when the variable can occasionally exhibit random jumps, such as oil prices or price of electricity (discrete exogenous event shocks can make prices jump up or down). Finally, these three stochastic processes can be mixed and matched as required.

The results on the right indicate the mean and standard deviation of all the iterations generated at each time step. If the Show All Iterations option is selected, each iteration pathway will be shown in a separate worksheet. The graph generated below shows a sample set of the iteration pathways.

Stochastic Process: Brownian Motion (Random Walk) with Drift

Start Value	100	Steps	100.00	Jump Rate	N/A
Drift Rate	5.00%	Iterations	10.00	Jump Size	N/A
Volatility	25.00%	Reversion Rate	N/A	Random Seed	1431155157
Horizon	10	Long-Term Value	N/A		

Time	Mean	Stdev
0.0000	100.00	0.00
0.1000	99.10	7.47
0.2000	96.03	7.22
0.3000	94.97	13.59
0.4000	97.39	15.57
0.5000	99.50	17.01
0.6000	97.79	20.92
0.7000	102.23	25.54
0.8000	106.54	26.54
0.9000	102.34	21.16
1.0000	102.77	20.86
1.1000	103.30	22.41
1.2000	103.27	19.23
1.3000	103.02	23.61
1.4000	97.78	19.65
1.5000	96.84	20.53
1.6000	100.92	25.22
1.7000	105.18	26.90
1.8000	100.75	30.33
1.9000	101.20	29.71
2.0000	103.67	36.95



2.1000	108.09	42.76
2.2000	111.58	42.61
2.3000	111.25	41.54
2.4000	108.47	35.22
2.5000	107.13	32.56
2.6000	108.95	32.95
2.7000	114.64	38.78
2.8000	114.13	36.61
2.9000	114.97	35.91
3.0000	114.33	39.90
3.1000	112.69	39.94
3.2000	115.11	39.89
3.3000	117.64	42.82
3.4000	114.70	39.91
3.5000	115.52	43.45
3.6000	117.60	49.89
3.7000	120.21	51.94
3.8000	116.64	53.52
3.9000	118.70	56.12
4.0000	113.19	56.71
4.1000	109.09	58.33
4.2000	103.70	52.23
4.3000	108.41	53.12
4.4000	108.67	56.30
4.5000	105.96	52.42
4.6000	106.12	55.80
4.7000	107.70	55.11
4.8000	109.43	58.43
4.9000	114.50	59.64
5.0000	110.44	53.91
5.1000	109.68	53.96
5.2000	113.38	62.37

Real Options Valuation

Risk Simulator
Hands-On Exercise

Nonlinear Extrapolation and Forecasting

This sample model is used to illustrate how to use Risk Simulator for:

1. Running Nonlinear Extrapolation and Forecasting
2. Comparing Extrapolation with Time-Series Forecasting

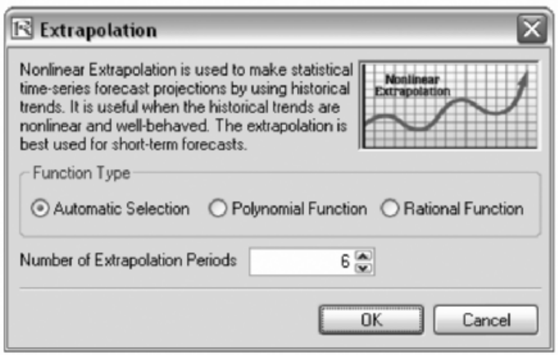
Model Background
File Name: Nonlinear Extrapolation.xls
This model provides some historical data on sales revenues for a firm. The goal of this exercise is to use Risk Simulator to run Nonlinear Extrapolation and forecast the revenues for the next several periods. The data are located in the Time-Series Data worksheet and are arranged by months, from Jan 2004 to Dec 2004. As the data are time-series in nature, we can apply extrapolation to forecast the results.

Nonlinear Extrapolation
Note that Nonlinear Extrapolation is involves making statistical projections by using historical trends that are projected for a specified period of time into the future. It is only used for time-series forecasts. Extrapolation is fairly reliable, relatively simple, and inexpensive. However, extrapolation, which assumes that recent and historical trends will continue, produces large forecast errors if discontinuities occur within the projected time period.

To run this model, simply:
1. Go to the Time-Series Data worksheet
2. Select the *Sales Revenue* data series (cells *H11:H22*) and select **Simulation | Forecasting | Nonlinear Extrapolation**
3. Extrapolate for *6 Periods* using the *Automatic Selection* option
4. Repeat the process by selecting the *Net Income* data series (cells *M11:M22*)
5. Select cells *H11:H22* again but this time run a Time-Series Analysis with forecast period of *6* with seasonality of *6*
6. Repeat Step 5's Time-Series Analysis on cells *M11:M22*
7. Compare the results from Extrapolation and Time-Series Analysis

Historical Sales Revenues
Polynomial Growth Rates

Year	Month	Period	Sales
2004	1	1	$1.00
2004	2	2	$6.73
2004	3	3	$20.52
2004	4	4	$45.25
2004	5	5	$83.59
2004	6	6	$138.01
2004	7	7	$210.87
2004	8	8	$304.44
2004	9	9	$420.89
2004	10	10	$562.34
2004	11	11	$730.85
2004	12	12	$928.43



Model Results Analysis

For your convenience, the analysis Report sheets are included. A fitted chart and forecast values are provided in the report as well as the error measures and a statistical summary of the methodology. Notice that when the historical data and future expectations are such that growth rates are nonlinear and are smooth, extrapolation works better (as in the case of the sales revenues, by both comparing the graphs visually and the corresponding RMSE values), but when seasonality occurs, time-series analysis works better (as in the case of net income).

Nonlinear Extrapolation

Statistical Summary

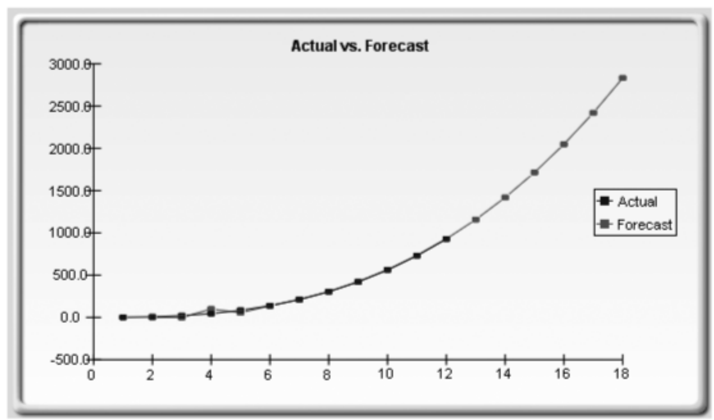
Extrapolation involves making statistical projections by using historical trends that are projected for a specified period of time into the future. It is only used for time-series forecasts. For cross-sectional or mixed panel data (time-series with cross-sectional data), multivariate regression is more appropriate. This methodology is useful when major changes are not expected, that is, causal factors are expected to remain constant or when the causal factors of a situation are not clearly understood. It also helps discourage introduction of personal biases into the process. Extrapolation is fairly reliable, relatively simple, and inexpensive. However, extrapolation, which assumes that recent and historical trends will continue, produces large forecast errors if discontinuities occur within the projected time period. That is, pure extrapolation of time series assumes that all we need to know is contained in the historical values of the series that is being forecasted. If we assume that past behavior is a good predictor of future behavior, extrapolation is appealing. This makes it a useful approach when all that is needed are many short-term forecasts.

This methodology estimates the f(x) function for any arbitrary x value, by interpolating a smooth nonlinear curve through all the x values, and using this smooth curve, extrapolates future x values beyond the historical data set. The methodology employs either the polynomial functional form or the rational functional form (a ratio of two polynomials). Typically, a polynomial functional form is sufficient for well-behaved data, however, rational functional forms are sometimes more accurate (especially with polar functions, i.e., functions with denominators approaching zero).

Period	Actual	Forecast Fit	Estimate Error
1	1.00		
2	6.73	1.00	
3	20.52	-1.42	-8.15
4	45.25	99.82	119.36
5	83.59	55.92	-46.67
6	138.01	136.71	14.39
7	210.87	211.96	1.69
8	304.44	304.43	-0.41
9	420.89	420.89	0.01
10	562.34	562.34	0.00
11	730.85	730.85	0.00
12	928.43	928.43	0.00
Forecast 13		1157.03	0.00
Forecast 14		1418.57	0.00
Forecast 15		1714.95	0.00
Forecast 16		2048.00	0.00
Forecast 17		2419.55	0.00
Forecast 18		2831.39	0.00

Error Measurements	
RMSE	19.6799
MSE	387.2974
MAD	10.2095
MAPE	31.56%
Theil's U	1.1210

Function Type: Rational



CHAPTER 9

Using the Past to Predict the Future

One of the more difficult tasks in risk analysis is forecasting, which includes the forecasting of any variable's future outcomes, for example, sales, revenues, machine failure rates, demand, costs, market share, competitive threats, and so forth. Recall from Chapter 8, Tomorrow's Forecast Today, that the most common quantitative or statistical approaches to forecasting include regression analysis, time-series analysis, nonlinear extrapolation, stochastic processes, and autoregressive integrated moving average (ARIMA). Time-series analysis, extrapolation, stochastics, and ARIMA are applicable for variables whose data are time-dependent, cross-sectional, or panel-based (both pooled time-dependent and cross-sectional data). Chapter 8 explores the basics of these methods and how to use Risk Simulator to forecast using these approaches, as well as some fundamental theories of these approaches. This chapter explores in more depth time-series and regression analysis through example computations. We start with time-series analysis by exploring the eight most common time-series methods or models as seen in Table 9.1. Regression analysis is then discussed, including the many pitfalls and dangers of applying regression analysis as a novice.

TIME-SERIES FORECASTING METHODOLOGY

Table 9.1 lists the eight most common time-series models, segregated by seasonality and trend. For instance, if the data variable has no trend or seasonality, then a single moving-average model or a single exponential-smoothing model would suffice. However, if seasonality exists but no discernible trend is present, either a seasonal additive or seasonal multiplicative model would be better, and so forth. The following sections explore these models in more detail through computational examples.

TABLE 9.1 The Eight Most Common Time-Series Methods

	No Seasonality	With Seasonality
No Trend	Single Moving Average Single Exponential Smoothing	Seasonal Additive Seasonal Multiplicative
With Trend	Double Moving Average Double Exponential Smoothing	Holt–Winters Additive Holt–Winters Multiplicative

NO TREND AND NO SEASONALITY

Single Moving Average

The single moving average is applicable when time-series data with no trend and seasonality exist. The approach simply uses an average of the actual historical data to project future outcomes. This average is applied consistently moving forward, hence the term *moving average*.

The value of the moving average (MA) for a specific length (n) is simply the summation of actual historical data (Y) arranged and indexed in time sequence (i).

$$MA_n = \frac{\sum_{i=1}^{n} Y_i}{n}$$

An example computation of a 3-month single moving average is seen in Figure 9.1. Here we see that there are 39 months of actual historical data and a 3-month moving average is computed.[1] Additional columns of calculations also exist in the example—calculations that are required to estimate the error of measurements in using this moving-average approach. These errors are important as they can be compared across multiple moving averages (i.e., 3-month, 4-month, 5-month, and so forth) as well as other time-series models (e.g., single moving average, seasonal additive model, and so forth) to find the best fit that minimizes these errors. Figures 9.2 to 9.4 show the exact calculations used in the moving-average model. Notice that the forecast-fit value in period 4 of 198.12 is a 3-month average of the prior three periods (months 1 through 3). The forecast-fit value for period 5 would then be the 3-month average of months 2 through 4. This process is repeated moving forward until month 40 (Figure 9.3), where every month after that, the forecast is fixed at 664.97. Clearly, this approach is not suitable if there is a trend (upward or downward over time) or if there is seasonality. Thus, error estimation is important when choosing the optimal time-series forecast model. Figure 9.2 illustrates a few additional columns of calculations required

Month	*Actual*	*Forecast Fit*	$\lvert Error \rvert$	$Error^2$	$\left\lvert \frac{Y_t - \hat{Y}_t}{Y_t} \right\rvert$	$\left[\frac{\hat{Y}_t - Y_t}{Y_{t-1}}\right]^2$	$\left[\frac{Y_t - Y_{t-1}}{Y_{t-1}}\right]$	*Error*	$\left[E_t - E_{t-1}\right]^2$
1	265.22	-	-	-	-	-	-	-	-
2	146.64	-	-	-	-	-	-	-	-
3	182.50		-	-	-	-	-	-	-
4	118.54	**198.12**	79.57	6332.12	67.13%	0.19	0.12	79.57	-
5	180.04	**149.23**	30.81	949.43	17.11%	0.07	0.27	-30.81	12185.39
6	167.45	**160.36**	7.09	50.20	4.23%	0.00	0.00	-7.09	562.99
7	231.75	**155.34**	76.41	5838.18	32.97%	0.21	0.15	-76.41	4805.61
8	223.71	**193.08**	30.63	938.22	13.69%	0.02	0.00	-30.63	2095.60
9	192.98	**207.64**	14.66	214.91	7.60%	0.00	0.02	14.66	2051.18
10	122.29	**216.15**	93.86	8808.84	76.75%	0.24	0.13	93.86	6271.97
11	336.65	**179.66**	157.00	24647.46	46.63%	1.65	3.07	-157.00	62925.98
12	186.50	**217.31**	30.81	949.17	16.52%	0.01	0.20	30.81	35270.22
13	194.27	**215.15**	20.88	435.92	10.75%	0.01	0.00	20.88	98.60
14	149.19	**239.14**	89.95	8091.27	60.29%	0.21	0.05	89.95	4771.05
15	210.06	**176.65**	33.41	1115.94	15.90%	0.05	0.17	-33.41	15216.99
16	272.91	**184.50**	88.40	7815.04	32.39%	0.18	0.09	-88.40	3024.67
17	191.93	**210.72**	18.79	352.98	9.79%	0.00	0.09	18.79	11489.77
18	286.94	**224.96**	61.97	3840.48	21.60%	0.10	0.25	-61.97	6522.06
19	226.76	**250.59**	23.83	567.99	10.51%	0.01	0.04	23.83	7362.34
20	303.38	**235.21**	68.17	4647.58	22.47%	0.09	0.11	-68.17	8465.03
21	289.72	**272.36**	17.36	301.32	5.99%	0.00	0.00	-17.36	2582.12
22	421.59	**273.29**	148.30	21993.55	35.18%	0.26	0.21	-148.30	17146.25
23	264.47	**338.23**	73.76	5440.32	27.89%	0.03	0.14	73.76	49310.98
24	342.30	**325.26**	17.04	290.41	4.98%	0.00	0.09	-17.04	8244.63
25	339.86	**342.79**	2.93	8.56	0.86%	0.00	0.00	2.93	398.71
26	439.90	**315.54**	124.35	15463.53	28.27%	0.13	0.09	-124.35	16199.87
27	315.54	**374.02**	58.48	3420.05	18.53%	0.02	0.08	58.48	33428.15
28	438.62	**365.10**	73.52	5404.80	16.76%	0.05	0.15	-73.52	17423.61
29	400.94	**398.02**	2.92	8.54	0.73%	0.00	0.01	-2.92	4983.77
30	437.37	**385.03**	52.34	2739.41	11.97%	0.02	0.01	-52.34	2442.13
31	575.77	**425.64**	150.13	22539.03	26.07%	0.12	0.10	-150.13	9563.01
32	407.33	**471.36**	64.03	4099.56	15.72%	0.01	0.09	64.03	45863.59
33	681.92	**473.49**	208.43	43442.59	30.57%	0.26	0.45	-208.43	74232.65
34	475.78	**555.01**	79.23	6277.13	16.65%	0.01	0.09	79.23	82746.68
35	581.17	**521.68**	59.49	3539.49	10.24%	0.02	0.05	-59.49	19243.79
36	647.82	**579.62**	68.20	4651.17	10.53%	0.01	0.01	-68.20	75.79
37	650.81	**568.26**	82.55	6814.39	12.68%	0.02	0.00	-82.55	205.92
38	677.54	**626.60**	50.94	2594.71	7.52%	0.01	0.00	-50.94	999.26
39	666.56	**658.72**	7.84	61.47	1.18%	0.00	0.00	-7.84	1857.46
Forecast 40	-	**664.97**	-	-	-	-	-	-	-
Forecast 41	-	**664.97**	-	-	-	-	-	-	-
Forecast 42	-	**664.97**	-	-	-	-	-	-	-

RMSE	79.00
MSE	6241.27
MAD	63.00
MAPE	20.80%
Thiel's U	0.80

$$MA_n = \frac{\sum_{i=1}^{n} Y_i}{n} \quad \forall\, i = 1,\dots, N$$

FIGURE 9.1 Single moving average (3 months).

for estimating the forecast errors. The values from these columns are used in Figure 9.4's error estimation.

Error Estimation (RMSE, MSE, MAD, MAPE, Theil's U)

Several different types of errors can be calculated for time-series forecast methods, including the mean-squared error (MSE), root mean-squared error

Month	*Actual*	*Forecast Fit*	$\lvert Error \rvert$	$Error^2$	$\left\lvert \frac{Y_t - \hat{Y}_t}{Y_t} \right\rvert$	$\left[\frac{\hat{Y}_t - Y_t}{Y_{t-1}}\right]^2$	$\left[\frac{Y_t - Y_{t-1}}{Y_{t-1}}\right]^2$	*Error*	$[E_t - E_{t-1}]^2$
1	265.22								
2	146.64								
3	182.50								
4	118.54	198.12	79.57	6332.12	67.13%	0.19	0.12	79.57	
								-30.81	12185.39

$$\frac{265.22 + 146.64 + 182.50}{3}$$

$$Abs\lvert 198.12 - 118.54 \rvert$$

$$79.57^2$$

$$\left\lvert \frac{118.54 - 198.12}{118.54} \right\rvert$$

$$\left[\frac{198.12 - 118.54}{182.50}\right]^2$$

$$\left[\frac{118.54 - 182.50}{182.50}\right]^2$$

$$198.12 - 118.54$$

$$[-30.81 - 79.57]^2$$

FIGURE 9.2 Calculating single moving average.

	35	581.17	**521.68**
	36	647.82	**579.62**
	37	650.81	**568.26**
	38	677.54	**626.60**
	39	666.56	**658.72**
Forecast	40	-	**664.97**
Forecast	41	-	**664.97**
Forecast	42	-	**664.97**

$$\frac{647.81 + 650.81 + 677.54}{3}$$

$$\frac{650.81 + 677.54 + 666.56}{3}$$

FIGURE 9.3 Forecasting with a single moving average.

RMSE	79.00
MSE	6241.27
MAD	63.00
MAPE	20.80%
Thiel's U	0.80

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(Error^2)_i}{n}} = \sqrt{MSE}$$

$$MSE = \sum_{i=1}^{n} \frac{(Error^2)_i}{n} = RMSE^2$$

$$MAD = \sum_{i=1}^{n} \frac{\lvert Error \rvert_i}{n}$$

$$MAPE = \sum_{i=1}^{n} \frac{\left\lvert \frac{Y_t - \hat{Y}_t}{Y_t} \right\rvert_i}{n}$$

$$Theil's\,U = \sqrt{\frac{\sum_{i=1}^{n}\left[\frac{\hat{Y}_t - Y_t}{Y_{t-1}}\right]_i^2}{\sum_{i=1}^{n}\left[\frac{Y_t - Y_{t-1}}{Y_{t-1}}\right]_i^2}}$$

FIGURE 9.4 Error estimation.

(RMSE), mean absolute deviation (MAD), and mean absolute percent error (MAPE).

The MSE is an absolute error measure that squares the errors (the difference between the actual historical data and the forecast-fitted data predicted by the model) to keep the positive and negative errors from canceling each other out. This measure also tends to exaggerate large errors by weighting the large errors more heavily than smaller errors by squaring them, which can help when comparing different time-series models. The MSE is calculated by simply taking the average of the *Error*^{2} column in Figure 9.1. RMSE is the square root of MSE and is the most popular error measure, also known as the *quadratic loss function*. RMSE can be defined as the average of the absolute values of the forecast errors and is highly appropriate when the cost of the forecast errors is proportional to the absolute size of the forecast error.

The MAD is an error statistic that averages the distance (absolute value of the difference between the actual historical data and the forecast-fitted data predicted by the model) between each pair of actual and fitted forecast data points. MAD is calculated by taking the average of the |*Error*| column in Figure 9.1, and is most appropriate when the cost of forecast errors is proportional to the absolute size of the forecast errors.

The MAPE is a relative error statistic measured as an average percent error of the historical data points and is most appropriate when the cost of the forecast error is more closely related to the percentage error than the numerical size of the error. This error estimate is calculated by taking the average of the

$$\left| \frac{Y_t - \hat{Y}_t}{Y_t} \right|$$

column in Figure 9.1, where Y_t is the historical data at time t, while $\hat{Y}_t$ is the fitted or predicted data point at time t using this time-series method. Finally, an associated measure is the Theil's U statistic, which measures the naivety of the model's forecast. That is, if the Theil's U statistic is less than 1.0, then the forecast method used provides an estimate that is statistically better than guessing. Figure 9.4 provides the mathematical details of each error estimate.

Single Exponential Smoothing

The second approach to use when no discernible trend or seasonality exists is the single exponential-smoothing method. This method weights past data with exponentially decreasing weights going into the past; that is, the more recent the data value, the greater its weight. This weighting largely overcomes the limitations of moving averages or percentage-change models. The weight used is termed the *alpha* measure. The method is illustrated in Figures 9.5 and 9.6 and uses the following model:

$$ESF_t = \alpha Y_{t-1} + (1 - \alpha)ESF_{t-1}$$

	Alpha		RMSE
	0.10		**126.26**
	Month	*Actual*	*Forecast Fit*
	1	265.22	
	2	146.64	265.22
	3	182.50	253.36
	4	118.54	246.28
	5	180.04	233.50
	6	167.45	228.16
	7	231.75	222.09
	8	223.71	223.05
	9	192.98	223.12
	10	122.29	220.10
	11	336.65	210.32
	12	186.50	222.96
	13	194.27	219.31
	14	149.19	216.81
	15	210.06	210.04
	16	272.91	210.05
	17	191.93	216.33
	18	286.94	213.89
	19	226.76	221.20
	20	303.38	221.75
	21	289.72	229.92
	22	421.59	235.90
	23	264.47	254.46
	24	342.30	255.47
	25	339.86	264.15
	26	439.90	271.72
	27	315.54	288.54
	28	438.62	291.24
	29	400.94	305.98
	30	437.37	315.47
	31	575.77	327.66
	32	407.33	352.47
	33	681.92	357.96
	34	475.78	390.35
	35	581.17	398.90
	36	647.82	417.12
	37	650.81	440.19
	38	677.54	461.26
	39	666.56	482.88
Forecast	40	-	501.25

$$ESF_t = \alpha\, Y_{t\text{-}1} + (1-\alpha)\, ESF_{t\text{-}1}$$

FIGURE 9.5 Single exponential smoothing.

Alpha
0.10

Month	*Actual*	*Forecast Fit*
1	265.22	
2	146.64	265.22
3	182.50	253.36
4	118.54	246.28
5	180.04	233.50
6	167.45	228.16
7	231.75	222.09
8	223.71	223.05

$\boxed{\hat{Y}_2 = Y_1 = 265.22}$

$\boxed{0.1(146.64) + (1 - 0.1)265.22}$

$$ESF_t = \alpha\, Y_{t\text{-}1} + (1-\alpha)\, ESF_{t\text{-}1}$$

FIGURE 9.6 Calculating single exponential smoothing.

where the exponential smoothing forecast (ESF_t) at time t is a weighted average between the actual value one period in the past (Y_{t-1}) and last period's forecast (ESF_{t-1}), weighted by the alpha parameter (α). Figure 9.6 shows an example of the computation. Notice that the first forecast-fitted value in month $2(\hat{Y}_2)$ is always the previous month's actual value (Y_1). The mathematical equation gets used only at month 3 or starting from the second forecast-fitted period.

Optimizing Forecasting Parameters

Clearly, in the single exponential-smoothing method, the alpha parameter was arbitrarily chosen as 0.10. In fact, the optimal alpha has to be obtained for the model to provide a good forecast. Using the model in Figure 9.5, Excel's Solver add-in package is used to find the optimal alpha parameter that minimizes the forecast errors. Figure 9.7 illustrates Excel's Solver add-in dialog box, where the target cell is set to the RMSE as the objective to be minimized by methodically changing the alpha parameter. As alpha should only be allowed to vary between 0.00 and 1.00 (because alpha is a weight given to the historical data and past period forecasts, and weights can never be less than zero or greater than one), additional constraints are also set up. The resulting optimal alpha value that minimizes forecast errors calculated

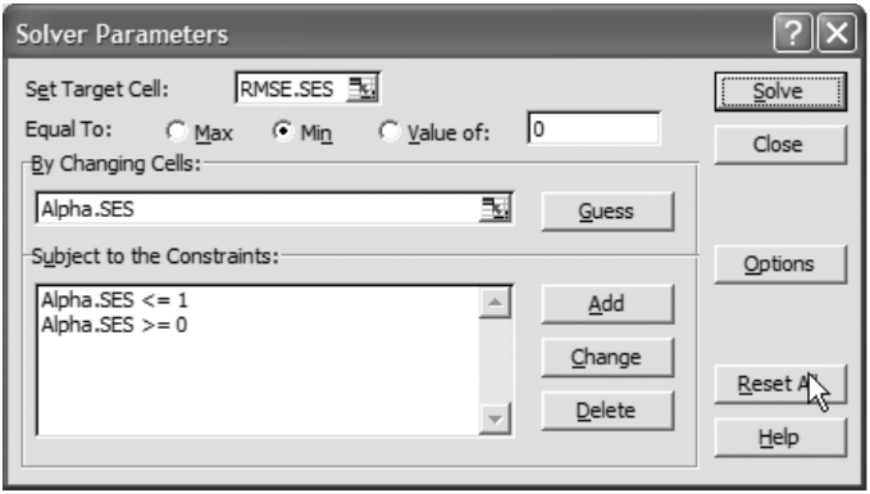


FIGURE 9.7 Optimizing parameters in single exponential smoothing.

by Solver is 0.4476. Therefore, entering this alpha value into the model will yield the best forecast values that minimize the errors.

WITH TREND BUT NO SEASONALITY

For data that exhibit a trend but no seasonality, the double moving-average and double exponential-smoothing methods work rather well.

Double Moving Average

The double moving-average method smoothes out past data by performing a moving average on a subset of data that represents a moving average of an original set of data. That is, a second moving average is performed on the first moving average. The second moving average application captures the trending effect of the data. Figures 9.8 and 9.9 illustrate the computation involved. The example shown is a 3-month double moving average and the forecast value obtained in period 40 is calculated using the following:

$$\text{Forecast} = 2MA_{1,t} - MA_{2,t} + \frac{2}{m-1}\left[MA_{1,t} - MA_{2,t}\right]$$

	Period	Actual	3-month MA_1	3-month MA_2	Forecast Fit
	1	265.22	-	-	-
	2	146.64	-	-	-
	3	182.50	-	-	-
	4	118.54	198.12	-	-
	5	180.04	149.23	-	-
	6	167.45	160.36	169.24	-
	7	231.75	155.34	154.98	142.61
	8	223.71	193.08	169.59	156.08
	9	192.98	207.64	185.35	240.05
	10	122.29	216.15	205.62	252.20
	11	336.65	179.66	201.15	237.20
	12	186.50	217.31	204.37	136.68
	13	194.27	215.15	204.04	243.18
	14	149.19	239.14	223.86	237.37
	15	210.06	176.65	210.31	269.69
	16	272.91	184.50	200.10	109.33
	17	191.93	210.72	190.62	153.32
	18	286.94	224.96	206.73	250.90
	19	226.76	250.59	228.76	261.44
	20	303.38	235.21	236.92	294.26
	21	289.72	272.36	252.72	231.78
	22	421.59	273.29	260.28	311.64
	23	264.47	338.23	294.62	299.29
	24	342.30	325.26	312.26	425.44
	25	339.86	342.79	335.42	351.26
	26	439.90	315.54	327.86	357.51
	27	315.54	374.02	344.12	290.91
	28	438.62	365.10	351.55	433.82
	29	400.94	398.02	379.04	392.19
	30	437.37	385.03	382.71	435.96
	31	575.77	425.64	402.90	389.66
	32	407.33	471.36	427.34	471.13
	33	681.92	473.49	456.83	559.39
	34	475.78	555.01	499.95	506.81
	35	581.17	521.68	516.72	665.12
	36	647.82	579.62	552.10	531.58
	37	650.81	568.26	556.52	634.66
	38	677.54	626.60	591.49	591.73
	39	666.56	658.72	617.86	696.81
Forecast	40	-	664.97	650.10	740.45

$$Forecast_{t+1} = 2MA_{1,t} - MA_{2,t} + \frac{2}{m-1}\left[MA_{1,t} - MA_{2,t}\right]$$

FIGURE 9.8 Double moving average (3 months).

Period	*Actual*	*3-month MA_1*	*3-month MA_2*	*Forecast Fit*
1	265.22			
2	146.64			
3	182.50			
4	118.54	198.12		
5	180.04	149.23		
6	167.45	160.36	169.24	
7	231.75	155.34	154.98	142.61
8	223.71	193.08	169.59	156.08

$$\frac{265.22 + 146.64 + 182.50}{3}$$

$$\frac{198.12 + 149.23 + 160.36}{3}$$

$$2(160.36) - 169.24 + \frac{2}{3-1}(160.36 - 169.24)$$

$$Forecast_{t+1} = 2MA_{1,t} - MA_{2,t} + \frac{2}{m-1}\left[MA_{1,t} - MA_{2,t}\right]$$

FIGURE 9.9 Calculating double moving average.

where the forecast value is twice the amount of the first moving average (MA_1) at time t, less the second moving average estimate (MA_2) plus the difference between the two moving averages multiplied by a correction factor (two divided into the number of months in the moving average, m, less one).

Double Exponential Smoothing

The second approach to use when the data exhibits a trend but no seasonality is the double exponential-smoothing method. Double exponential smoothing applies single exponential smoothing twice, once to the original data and then to the resulting single exponential-smoothing data. An alpha (α) weighting parameter is used on the first or single exponential smoothing (*SES*) while a beta (β) weighting parameter is used on the second or double exponential smoothing (*DES*). This approach is useful when the historical data series is not stationary. Figure 9.10 illustrates the double exponential-smoothing model, while Figure 9.11 shows Excel's Solver add-in dialog box used to find the optimal alpha and beta parameters that minimize the forecast errors. Figure 9.12 shows the computational details. The forecast is calculated using the following:

$$DES_t = \beta\left(SES_t - SES_{t-1}\right) + \left(1 - \beta\right)DES_{t-1}$$
$$SES_t = \alpha Y_t + \left(1 - \alpha\right)\left(SES_{t-1} + DES_{t-1}\right)$$

Note that the starting value (period 1 for DES in Figure 9.10) can take on different values other than the one shown. In some instances, zero is used when no prior information is available.

		Alpha 0.1593	Beta 0.3919	RMSE 70.81
Period	*Actual*	*SES*	*DES*	*Forecast Fit*
1	265.22	265.22	0.00	-
2	146.64	246.33	–7.40	-
3	182.50	229.94	–10.93	238.93
4	118.54	203.01	–17.20	219.01
5	180.04	184.89	–17.56	185.81
6	167.45	167.35	–17.55	167.33
7	231.75	162.85	–12.44	149.80
8	223.71	162.09	–12.44	150.42
9	192.98	160.41	–5.44	154.23
10	122.29	149.76	–7.48	154.96
11	336.65	173.24	4.65	142.28
12	186.50	179.27	5.19	177.90
13	194.27	186.02	5.80	184.46
14	149.19	185.03	3.14	191.82
15	210.06	191.66	4.51	188.17
16	272.91	208.39	9.30	196.17
17	191.93	213.59	7.69	217.69
18	286.94	231.74	11.79	221.28
19	226.76	240.86	10.74	243.53
20	303.38	259.85	13.98	251.60
21	289.72	276.35	14.97	273.82
22	421.59	312.07	23.10	291.32
23	264.47	323.91	18.69	335.17
24	342.30	342.55	18.67	342.60
25	339.86	357.82	17.33	361.22
26	439.90	385.46	21.38	375.15
27	315.54	392.30	15.68	406.84
28	438.62	412.85	17.59	407.97
29	400.94	425.74	15.75	430.44
30	437.37	440.83	15.49	441.49
31	575.77	475.35	22.95	456.32
32	407.33	483.81	17.27	498.30
33	681.92	529.88	28.56	501.08
34	475.78	545.27	23.40	558.44
35	581.17	570.66	24.18	568.67
36	647.82	603.28	27.49	594.84
37	650.81	633.96	28.74	630.77
38	677.54	665.06	29.66	662.69
39	666.56	690.24	27.91	694.72
Forecast 40	-	-	-	718.14
Forecast 41	-	-	-	746.05
Forecast 42	-	-	-	773.95
Forecast 43	-	-	-	801.86

$$DES_t = \beta\ (SES_t - SES_{t-1}) + (1-\beta)\ DES_{t-1}$$
$$SES_t = \alpha\ Y_t + (1-\alpha)(SES_{t-1} + DES_{t-1})$$

FIGURE 9.10 Double exponential smoothing.

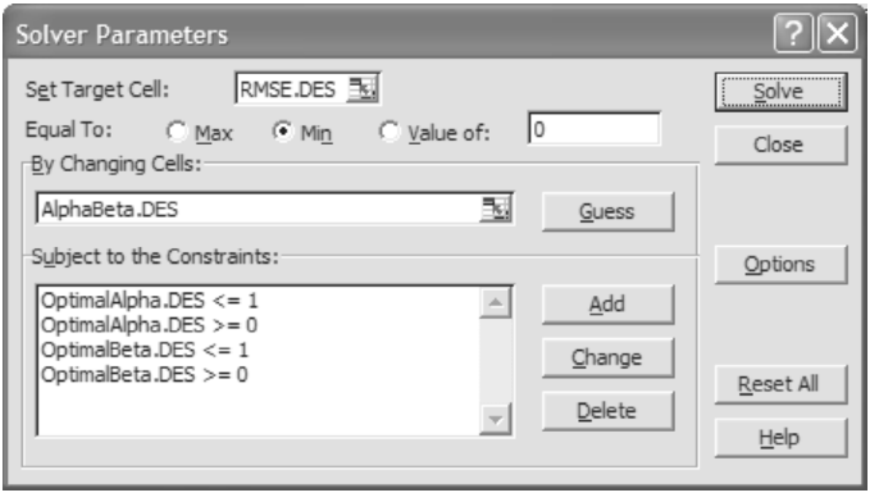
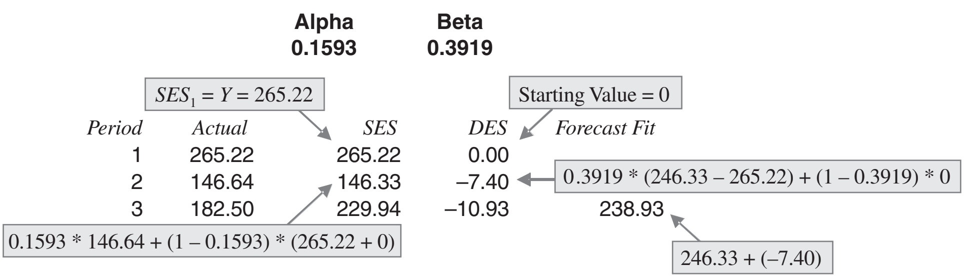


FIGURE 9.11 Optimizing parameters in double exponential smoothing.



$$DES_t = \beta\,(SES_t - SES_{t-1}) + (1-\beta)\,DES_{t-1}$$
$$SES_t = \alpha\,Y_t + (1-\alpha)\,(SES_{t-1} + DES_{t-1})$$

FIGURE 9.12 Calculating double exponential smoothing.

NO TREND BUT WITH SEASONALITY

Additive Seasonality

If the time-series data has no appreciable trend but exhibits seasonality, then the additive seasonality and multiplicative seasonality methods apply. The additive seasonality method is illustrated in Figures 9.13 and 9.14. The

	Level Alpha	Seasonal Gamma		RMSE
	0.33	0.40		93.54

Period	*Actual*	*Level*	*Seasonality*	*Forecast Fit*
1	265.22	-	87.00	-
2	146.64	-	-31.59	-
3	182.50	-	4.27	-
4	118.54	178.23	-59.68	-
5	180.04	150.44	63.85	265.22
6	167.45	166.29	-18.38	118.86
7	231.75	186.25	20.90	170.56
8	223.71	217.93	-33.28	126.57
9	192.98	188.97	39.72	281.78
10	122.29	173.22	-31.51	170.58
11	336.65	219.70	59.63	194.12
12	186.50	219.73	-33.26	186.42
13	194.27	198.47	22.01	259.45
14	149.19	192.67	-36.34	166.96
15	210.06	178.90	48.15	252.31
16	272.91	220.40	1.32	145.63
17	191.93	203.94	8.29	242.41
18	286.94	242.86	-3.91	167.60
19	226.76	221.90	30.69	291.01
20	303.38	248.05	23.10	223.23
21	289.72	258.93	17.36	256.34
22	421.59	313.26	41.35	255.02
23	264.47	287.34	9.09	343.95
24	342.30	297.73	31.76	310.44
25	339.86	305.81	24.09	315.09
26	439.90	336.05	66.55	347.16
27	315.54	326.40	1.05	345.15
28	438.62	352.64	53.62	358.16
29	400.94	360.53	30.67	376.73
30	437.37	363.89	69.35	427.08
31	575.77	432.65	58.34	364.94
32	407.33	406.90	32.17	486.27
33	681.92	486.59	97.07	437.57
34	475.78	460.45	47.56	555.94
35	581.17	480.80	75.29	518.79
36	647.82	524.78	68.82	512.97
37	650.81	534.22	104.94	621.84
38	677.54	565.45	73.58	581.79
39	666.56	573.87	82.31	640.74

$$Level\ L_t = \alpha(Y_t - S_{t-s}) + (1-\alpha)(L_{t-1})$$
$$Seasonality\ S_t = \gamma(Y_t - L_t) + (1-\gamma)(S_{t-s})$$
$$Forecast\ \ F_{t+m} = L_t + S_{t+m-s}$$

FIGURE 9.13 Additive seasonality with no trend.

Level Alpha	Seasonal Gamma	RMSE
0.33	0.40	85.18

Period	*Actual*	*Level*	*Seasonality*	*Forecast Fit*
1	265.22		87.00	
2	146.64		-31.59	
3	182.50		4.27	
4	118.54	178.23	-59.68	
5	180.04	150.44	63.85	265.22

$\boxed{Y_t - 178.23}$

$\boxed{178.23 + 87.00}$

$\boxed{\dfrac{265.22 + 146.64 + 182.50 + 118.54}{4}}$

$\boxed{0.40(180.04 - 150.44) + (1 - 0.40)(87.00)}$

$\boxed{0.33(180.04 - 87.00) + (1 - 0.33)(178.23)}$

$$\textit{Level } L_t = \alpha(Y_t - S_{t-s}) + (1-\alpha)(L_{t-1})$$
$$\textit{Seasonality } S_t = \gamma(Y_t - L_t) + (1-\gamma)(S_{t-s})$$
$$\textit{Forecast } F_{t+m} = L_t + S_{t+m-s}$$

FIGURE 9.14 Calculating seasonal additive.

additive seasonality model breaks the historical data into a level (L) or base case component as measured by the alpha parameter (α), and a seasonality (S) component measured by the gamma parameter (γ). The resulting forecast value is simply the addition of this base case level to the seasonality value. Quarterly seasonality is assumed in the example. (Note that calculations are rounded.)

Multiplicative Seasonality

Similarly, the multiplicative seasonality model requires the alpha and gamma parameters. The difference from additive seasonality is that the model is multiplicative, for example, the forecast value is the multiplication between the base case level and seasonality factor. Figures 9.15 and 9.16 illustrate the computations required. Quarterly seasonality is assumed in the example. (Calculations are rounded.)

		Level Alpha 0.22	Seasonal Gamma 0.64	RMSE 95.65
Period	*Actual*	*Level*	*Seasonality*	*Forecast Fit*
1	265.22	-	1.49	-
2	146.64	-	0.82	-
3	182.50	-	1.02	-
4	118.54	178.23	0.67	-
5	180.04	165.35	1.23	265.22
6	167.45	173.93	0.91	136.04
7	231.75	185.72	1.17	178.11
8	223.71	219.61	0.89	123.53
9	192.98	205.42	1.04	270.67
10	122.29	189.36	0.74	187.42
11	336.65	211.65	1.44	221.04
12	186.50	211.10	0.89	188.67
13	194.27	205.43	0.98	220.57
14	149.19	204.47	0.73	152.37
15	210.06	191.32	1.22	294.08
16	272.91	217.55	1.12	169.58
17	191.93	212.61	0.93	213.50
18	286.94	252.73	0.99	156.05
19	226.76	237.67	1.05	308.43
20	303.38	245.03	1.20	266.66
21	289.72	259.92	1.05	228.13
22	421.59	297.16	1.26	257.56
23	264.47	286.97	0.97	311.99
24	342.30	286.78	1.19	343.32
25	339.86	295.18	1.11	300.72
26	439.90	307.02	1.37	373.34
27	315.54	311.30	1.00	297.12
28	438.62	323.87	1.30	371.87
29	400.94	331.95	1.17	360.91
30	437.37	328.97	1.34	455.55
31	575.77	384.87	1.32	328.02
32	407.33	368.95	1.17	499.11
33	681.92	416.60	1.47	433.22
34	475.78	402.47	1.24	560.30
35	581.17	411.24	1.38	529.84
36	647.82	442.93	1.36	482.55
37	650.81	442.86	1.47	651.26
38	677.54	466.08	1.38	549.47
39	666.56	470.02	1.40	642.45

$$\text{Level } L_t = \alpha(Y_t / S_{t-s}) + (1-\alpha)(L_{t-1})$$
$$\text{Seasonality } S_t = \gamma(Y_t / L_t) + (1-\gamma)(S_{t-s})$$
$$\text{Forecast } F_{t+m} = L_t S_{t+m-s}$$

FIGURE 9.15 Multiplicative seasonality with no trend.

		Level	Seasonal	RMSE
		Alpha	Gamma	85.18
		0.22	0.64	

Period	*Actual*	*Level*	*Seasonality*	*Forecast Fit*
1	265.22		1.49	
2	146.64		0.82	
3	182.50		1.02	
4	118.54	178.23	0.67	
5	180.04	165.35	1.23	265.22

$$\frac{Y_t}{178.23}$$

$$178.23(1.49)$$

$$\frac{265.22+146.64+182.50+118.54}{4}$$

$$0.22(180.04/1.49)+(1-0.22)(178.23)$$

$$0.64(180.04/165.35)+(1-0.64)(1.49)$$

$$\text{Level } L_t = \alpha(Y_t / S_{t-s}) + (1-\alpha)(L_{t-1})$$
$$\text{Seasonality } S_t = \gamma(Y_t / L_t) + (1-\gamma)(S_{t-s})$$
$$\text{Forecast } F_{t+m} = L_t S_{t+m-s}$$

FIGURE 9.16 Calculating seasonal multiplicative.

WITH SEASONALITY AND WITH TREND

When both seasonality and trend exist, more advanced models are required to decompose the data into their base elements: a base case level (L) weighted by the alpha parameter (α); a trend component (b) weighted by the beta parameter (β); and a seasonality component (S) weighted by the gamma parameter (γ). Several methods exist, but the two most common are the Holt–Winters additive seasonality and Holt–Winters multiplicative seasonality methods.

Holt–Winters Additive Seasonality

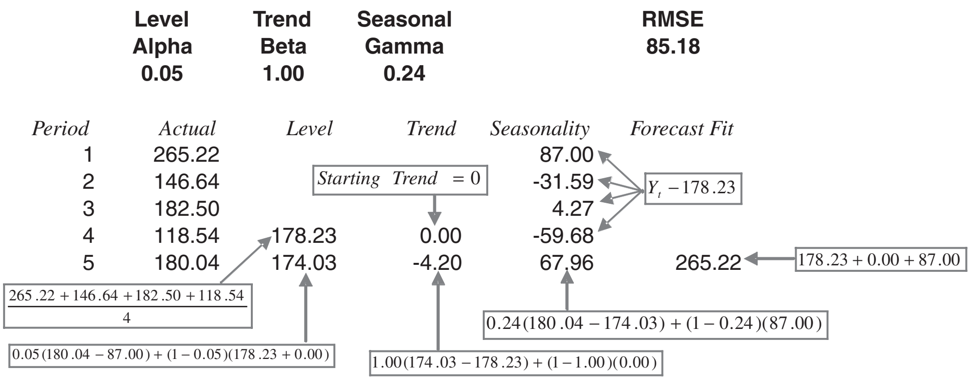
Figures 9.17 and 9.18 illustrate the required computations for determining a Holt–Winters additive forecast model. (Calculations are rounded.)

	Level	Trend	Seasonal		RMSE
	Alpha	**Beta**	**Gamma**		**77.03**
	0.05	**1.00**	**0.24**		

Period	*Actual*	*Level*	*Trend*	*Seasonality*	*Forecast Fit*
1	265.22	-	-	87.00	-
2	146.64	-	-	-31.59	-
3	182.50	-	-	4.27	-
4	118.54	178.23	0.00	-59.68	-
5	180.04	174.03	-4.20	67.96	265.22
6	167.45	171.27	-2.76	-25.06	138.25
7	231.75	171.42	0.15	17.45	172.79
8	223.71	177.07	5.65	-34.69	111.89
9	192.98	179.89	2.81	55.06	250.69
10	122.29	180.96	1.07	-32.96	157.64
11	336.65	188.78	7.83	48.11	199.48
12	186.50	197.82	9.04	-29.20	161.92
13	194.27	203.53	5.71	39.94	261.92
14	149.19	207.90	4.37	-39.01	176.27
15	210.06	209.79	1.89	36.86	260.38
16	272.91	216.14	6.35	-8.99	182.49
17	191.93	219.01	2.87	24.19	262.43
18	286.94	227.01	8.00	-15.76	182.87
19	226.76	232.79	5.78	26.78	271.87
20	303.38	242.20	9.41	7.50	229.58
21	289.72	252.30	10.10	27.30	275.80
22	421.59	271.02	18.71	23.34	246.64
23	264.47	287.17	16.15	15.15	316.51
24	342.30	304.87	17.70	14.54	310.82
25	339.86	322.08	17.21	25.06	349.87
26	439.90	343.09	21.01	40.61	362.63
27	315.54	360.97	17.88	0.91	379.26
28	438.62	381.07	20.10	24.65	393.38
29	400.94	399.93	18.86	19.41	426.24
30	437.37	417.70	17.77	35.69	459.40
31	575.77	442.34	24.64	32.06	436.38
32	407.33	462.83	20.49	5.81	491.63
33	681.92	492.14	29.31	59.45	502.72
34	475.78	517.45	25.31	17.50	557.14
35	581.17	543.06	25.62	33.48	574.81
36	647.82	572.29	29.23	22.20	574.49
37	650.81	601.02	28.73	57.18	660.98
38	677.54	631.24	30.22	24.27	647.26
39	666.56	660.07	28.82	27.14	694.95

$$\text{Level } L_t = \alpha(Y_t - S_{t-s}) + (1-\alpha)(L_{t-1} + b_{t-1})$$
$$\text{Trend } b_t = \beta(L_t - L_{t-1}) + (1-\beta)(b_{t-1})$$
$$\text{Seasonality } S_t = \gamma(Y_t - L_t) + (1-\gamma)(S_{t-s})$$
$$\text{Forecast } F_{t+m} = L_t + mb_t + S_{t+m-s}$$

FIGURE 9.17 Holt–Winters additive seasonality with trend.



$$\textit{Level}\ L_t = \alpha(Y_t - S_{t-s}) + (1-\alpha)(L_{t-1} + b_{t-1})$$
$$\textit{Trend}\ b_t = \beta(L_t - L_{t-1}) + (1-\beta)(b_{t-1})$$
$$\textit{Seasonality}\ S_t = \gamma(Y_t - L_t) + (1-\gamma)(S_{t-s})$$
$$\textit{Forecast}\ F_{t+m} = L_t + mb_t + S_{t+m-s}$$

FIGURE 9.18 Calculating Holt–Winters additive.

Holt–Winters Multiplicative Seasonality

Figures 9.19 and 9.20 show the required computation for determining a Holt–Winters multiplicative forecast model when both trend and seasonality exist. (Calculations are rounded.)

REGRESSION ANALYSIS

This section deals with using regression analysis for forecasting purposes. It is assumed that the reader is sufficiently knowledgeable about the fundamentals of regression analysis. Instead of focusing on the detailed theoretical mechanics of the regression equation, we look at the basics of applying regression analysis and work through the various relationships that a regression analysis can capture, as well as the common pitfalls in regression, including the problems of outliers, nonlinearities, heteroskedasticity, autocorrelation, and structural breaks.

The general bivariate linear regression equation takes the form of

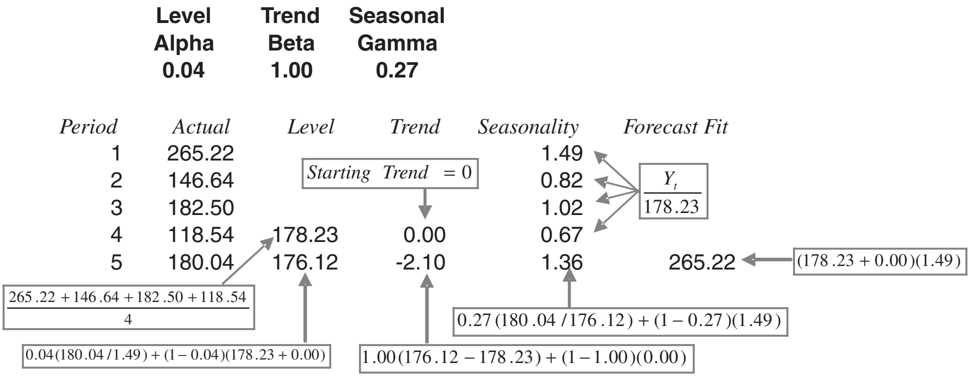
$$Y = \beta_0 + \beta_1 X + \varepsilon$$

	Level	Trend	Seasonal		RMSE
	Alpha	**Beta**	**Gamma**		**79.15**
	0.04	**1.00**	**0.27**		

Period	*Actual*	*Level*	*Trend*	*Seasonality*	*Forecast Fit*
1	265.22	-	-	1.49	-
2	146.64	-	-	0.82	-
3	182.50	-	-	1.02	-
4	118.54	178.23	0.00	0.67	-
5	180.04	176.12	-2.10	1.36	265.22
6	167.45	175.11	-1.02	0.86	143.18
7	231.75	176.01	0.90	1.10	178.26
8	223.71	182.75	6.75	0.82	117.67
9	192.98	187.75	5.00	1.27	257.93
10	122.29	190.90	3.15	0.80	165.60
11	336.65	198.12	7.22	1.27	214.19
12	186.50	206.17	8.06	0.84	167.87
13	194.27	211.98	5.81	1.17	272.12
14	149.19	216.64	4.66	0.77	174.13
15	210.06	219.27	2.63	1.18	280.20
16	272.91	225.66	6.39	0.94	186.67
17	191.93	229.53	3.88	1.08	272.38
18	286.94	238.53	9.00	0.89	179.57
19	226.76	245.48	6.95	1.11	292.61
20	303.38	254.99	9.51	1.01	237.70
21	289.72	264.63	9.63	1.09	286.13
22	421.59	281.63	17.00	1.05	243.42
23	264.47	296.40	14.77	1.05	331.98
24	342.30	312.20	15.80	1.03	314.05
25	339.86	327.45	15.25	1.07	355.98
26	439.90	345.45	18.00	1.11	361.10
27	315.54	361.12	15.67	1.00	382.29
28	438.62	378.54	17.42	1.07	389.23
29	400.94	395.15	16.61	1.06	424.62
30	437.37	411.07	15.91	1.10	458.54
31	575.77	432.37	21.30	1.09	428.40
32	407.33	451.03	18.66	1.02	484.20
33	681.92	476.14	25.11	1.16	496.30
34	475.78	498.73	22.59	1.06	551.41
35	581.17	521.70	22.97	1.10	569.70
36	647.82	547.93	26.23	1.07	556.94
37	650.81	573.70	25.77	1.15	665.46
38	677.54	600.92	27.22	1.08	635.58
39	666.56	627.35	26.43	1.09	690.07

$$Level\ L_t = \alpha(Y_t / S_{t-s}) + (1-\alpha)(L_{t-1} + b_{t-1})$$
$$Trend\ b_t = \beta(L_t - L_{t-1}) + (1-\beta)(b_{t-1})$$
$$Seasonality\ S_t = \gamma(Y_t / L_t) + (1-\gamma)(S_{t-s})$$
$$Forecast\ F_{t+m} = (L_t + mb_t)S_{t+m-s}$$

FIGURE 9.19 Holt–Winters multiplicative seasonality with trend.



$$\textit{Level } L_t = \alpha(Y_t / S_{t-s}) + (1-\alpha)(L_{t-1} + b_{t-1})$$
$$\textit{Trend } b_t = \beta(L_t - L_{t-1}) + (1-\beta)(b_{t-1})$$
$$\textit{Seasonality } S_t = \gamma(Y_t / L_t) + (1-\gamma)(S_{t-s})$$
$$\textit{Forecast } F_{t+m} = (L_t + mb_t)S_{t+m-s}$$

FIGURE 9.20 Calculating Holt–Winters multiplicative.

where β_0 is the intercept, β_1 is the slope, and ε is the error term. It is bivariate as there are only two variables, a Y or dependent variable, and an X or independent variable, where X is also known as the regressor (sometimes a bivariate regression is also known as a univariate regression as there is only a single independent variable X). The dependent variable is named as such as it *depends* on the independent variable, for example, sales revenue depends on the amount of marketing costs expended on a product's advertising and promotion, making the dependent variable sales and the independent variable marketing costs. An example of a bivariate regression is seen as simply inserting the best-fitting line through a set of data points in a two-dimensional plane as seen on the left panel in Figure 9.21. In other cases, a multivariate regression can be performed, where there are multiple or n number of independent X variables, where the general regression equation will now take the form of $Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 \ldots + \beta_nX_n + \varepsilon$. In this case, the best-fitting line will be within an $n + 1$ dimensional plane.

However, fitting a line through a set of data points in a scatter plot as in Figure 9.21 may result in numerous possible lines. The best-fitting line is defined as the single unique line that minimizes the total vertical errors, that is, the sum of the absolute distances between the actual data points (Y_i) and

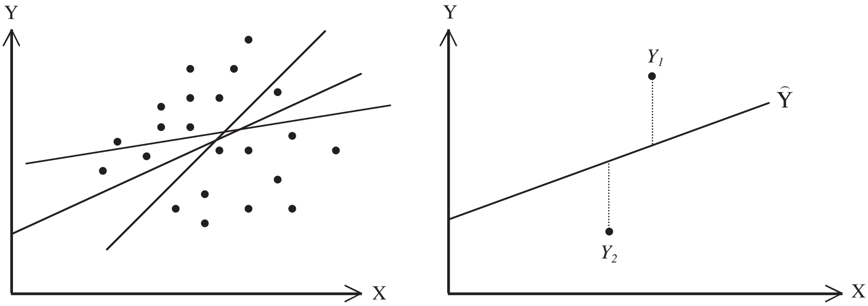


FIGURE 9.21 Bivariate regression.

the estimated line ($\hat{Y}$) as shown on the right panel of Figure 9.21. In order to find the best-fitting line that minimizes the errors, a more sophisticated approach is required, that is, regression analysis. Regression analysis therefore finds the unique best-fitting line by requiring that the total errors be minimized, or by calculating

$$\text{Min} \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2$$

where only one unique line minimizes this sum of squared errors. The errors (vertical distance between the actual data and the predicted line) are squared to avoid the negative errors from canceling out the positive errors. Solving this minimization problem with respect to the slope and intercept requires calculating a first derivative and setting them equal to zero:

$$\frac{d}{d\beta_0} \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 = 0 \qquad \text{and} \qquad \frac{d}{d\beta_1} \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 = 0$$

which yields the *least squares regression equations* seen in Figure 9.22.

See Appendix B—Ordinary Least Squares at the end of this chapter for more details on optimizing this line to find the best-fitting line.

Example Given the following sales amounts ($ millions) and advertising sizes (measured as linear inches by summing up all the sides of an ad) for a local newspaper, answer the accompanying questions.

Advertising size (inch)	**12**	**18**	**24**	**30**	**36**	**42**	**48**
Sales ($ millions)	5.9	5.6	5.5	7.2	8.0	7.7	8.4

$$\beta_1 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{n}(X_i - \overline{X})^2} = \frac{\sum_{i=1}^{n} X_i Y_i - \frac{\sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i}{n}}{\sum_{i=1}^{n} X_i^2 - \frac{\left(\sum_{i=1}^{n} X_i\right)^2}{n}}$$

$$\text{and} \quad \beta_0 = \overline{Y} - \beta_1 \overline{X}$$

FIGURE 9.22 Least squares regression equations.

1. Which is the dependent variable and which is the independent variable? *The independent variable is advertising size, whereas the dependent variable is sales.*
2. Manually calculate the slope (β_1) and the intercept (β_0) terms.

X	***Y***	***XY***	***X*²**	***Y*²**
12	5.9	70.8	144	34.81
18	5.6	100.8	324	31.36
24	5.5	132.0	576	30.25
30	7.2	216.0	900	51.84
36	8.0	288.0	1296	64.00
42	7.7	323.4	1764	59.29
48	8.4	403.2	2304	70.56
$\Sigma(X) = 210$	$\Sigma(Y) = 48.3$	$\Sigma(XY) = 1534.2$	$\Sigma(X^2) = 7308$	$\Sigma(Y^2) = 342.11$

$$\beta_1 = \frac{1534.2 - \frac{210(48.3)}{7}}{7308 - \frac{210^2}{7}} = 0.0845 \quad and \quad \beta_0 = \frac{48.3}{7} - 0.0845\left[\frac{210}{7}\right] = 4.3643$$

3. What is the estimated regression equation?

$$Y = 4.3643 + 0.0845X \quad or \quad Sales = 4.3643 + 0.0845(\text{Size})$$

4. What would the level of sales be if we purchase a 28-inch ad?

$$Y = 4.3643 + 0.0845\ (28) = \$6.73 \text{ million in sales}$$

(Note that we only predict or forecast and cannot say for certain. This is only an expected value or on average.)

Regression Output

Using the data in the previous example, a regression analysis can be performed either using Excel's Data Analysis add-in or Risk Simulator software.[2] Figure 9.23 shows Excel's regression analysis output. Notice that the coefficients on the intercept and *X* variable confirm the results we obtained in the manual calculation.

The same regression analysis can be performed using Risk Simulator.[3] The results obtained through Risk Simulator are seen in Figure 9.24. Notice again the identical answers to the slope and intercept calculations. Clearly, there are significant amounts of additional information obtained through the Excel and Risk Simulator analyses. Most of these additional statistical outputs pertain to goodness-of-fit measures, that is, a measure of how accurate and statistically reliable the model is.

Goodness-of-Fit

Goodness-of-fit statistics provide a glimpse into the accuracy and reliability of the estimated regression model. They usually take the form of a t-statistic, F-statistic, R-squared statistic, adjusted R-squared statistic, Durbin–Watson statistic, and their respective probabilities. (See the t-statistic, F-statistic, and critical Durbin–Watson tables at the end of this book for the corresponding critical values used later in this chapter.) The following sections discuss some of the more common regression statistics and their interpretation.

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.9026
R Square	0.8146
Adjusted R Square	0.7776
Standard Error	0.5725
Observations	7

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	7.2014	7.2014	21.9747	0.0054
Residual	5	1.6386	0.3277		
Total	6	8.8400			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	4.3643	0.5826	7.4911	0.0007	2.8667	5.8619	2.8667	5.8619
X Variable 1	0.0845	0.0180	4.6877	0.0054	0.0382	0.1309	0.0382	0.1309

FIGURE 9.23 Regression output from Excel's Data Analysis add-in.

Regression Analysis Report

Regression Statistics

R-Squared (Coefficient of Determination)	0.8146
Adjusted R-Squared	0.7776
Multiple R (Multiple Correlation Coefficient)	0.9026
Standard Error of the Estimates (SEy)	0.5725
nObservations	7

The R-squared or coefficient of determination indicates that 0.81 of the variation in the dependent variable can be explained and accounted for by the independent variables in this regression analysis. However, in a multiple regression, the adjusted R-squared takes into account the existence of additional independent variables or regressors and adjusts this R-squared value to a more accurate view of the regression's explanatory power. Hence, only 0.78 of the variation in the dependent variable can be explained by the regressors.

The multiple correlation coefficient (Multiple R) measures the correlation between the actual dependent variable (Y) and the estimated or fitted (Y) based on the regression equation. This is also the square root of the coefficient of determination (R-Squared.)

The standard error of the estimates (SE_y) describes the dispersion of data points above and below the regression line or plane. This value is used as part of the calculation to obtain the confidence interval of the estimates later.

Regression Results

	Intercept	Ad Size
Coefficients	4.3643	0.0845
Standard Error	0.5826	0.0180
t-Statistic	7.4911	4.6877
p-Value	0.0007	0.0054
Lower 5%	2.8667	0.0382
Upper 95%	5.8619	0.1309

Degrees of Freedom	
Degrees of Freedom for Regression	1
Degrees of Freedom for Residual	5
Total Degrees of Freedom	6

Hypothesis Test	
Critical t-Statistic (99% confidence with df of 5)	4.0321
Critical t-Statistic (95% confidence with df of 5)	2.5706
Critical t-Statistic (90% confidence with df of 5)	2.0150

The coefficients provide the estimated regression intercept and slopes. For instance, the coefficients are estimates of the true population b values in the following regression equation: $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_2 X_2$. The standard error measures how accurate the predicted Coefficients are, and the t-Statistics are the ratios of each predicted coefficient to its standard error.

The t-statistic is used in hypothesis testing, where we set the null hypothesis (H_0) such that the real mean of the coefficient = 0, and the alternate hypothesis (H_a) such that the real mean of the Coefficient is not equal to 0. A t-test is performed and the calculated t-statistic is compared to the critical values at the relevant degrees of freedom for residual. The t-test is very important as it calculates if each of the coefficients is statistically significant in the presence of the other regressors. This means that the t-test statistically verifies whether a regressor or independent variable should remain in the regression or it should be dropped.

The coefficient is statistically significant if its calculated t-statistic exceeds the critical t-statistic at the relevant degrees of freedom (df). The three main confidence levels used to test for significance are 90%, 95%, and 99%. If a coefficient's t-statistic exceeds the critical level, it is considered statistically significant. Alternatively, the p-value calculates each t-statistic's probability of occurrence, which means that the smaller the p-value, the more significant the coefficient. The usual significant levels for the p-value are 0.01, 0.05, and 0.10, corresponding to the 99%, 95%, and 99% confidence levels.

The coefficients with their p-values highlighted in blue indicate that they are statistically significant at the 95% confidence or 0.05 alpha level, while those highlighted in red indicate that they are not statistically significant at any of the alpha levels.

FIGURE 9.24 Regression output from Risk Simulator software.

The R-squared (R^2), or coefficient of determination, is an error measurement that looks at the percent variation of the dependent variable that can be explained by the variation in the independent variable for a regression analysis. The coefficient of determination can be calculated by:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} = 1 - \frac{SSE}{TSS}$$

Analysis of Variance

	Sums of Squares	Mean of Squares	F-Statistic	P-Value	Hypothesis Test	
Regression	7.2014	7.2014	21.9747	0.0054	Critical t-Statistic (99% confidence with df of 4 and 3)	16.2582
Residual	1.6386	0.3277			Critical t-Statistic (95% confidence with df of 4 and 3)	6.6079
Total	8.8400				Critical t-Statistic (90% confidence with df of 4 and 3)	4.0604

The analysis of variance (ANOVA) table provides an F-test of the regression model's overall statistical significance. Instead of looking at individual regressors as in the t-test, the F-test looks at all the estimated Coefficients statistical properties. The F-statistic is calculated as the ratio of the regression's mean of squares to the residual's mean of squares. The numerator measures how much of the regression is explained, while the denominator measures how much is unexplained. Hence, the larger the F-statistic, the more significant the model. The corresponding p-value is calculated to test the null hypothesis (H_0) where all the coefficients are simultaneously equal to zero, versus the alternate hypothesis (H_a) that they are all simultaneously different from zero, indicating a significant overall regression model. If the p-value is smaller than the 0.01, 0.05, or 0.10 alpha significance, then the regression is significant. The same approach can be applied to the F-statistic by comparing the calculated F-statistic with the critical F-values at various significance levels.

Forecasting

Period	Actual (Y)	Forecast (F)	Error (E)
1	5.9	5.3786	0.5214
2	5.6	5.8857	(0.2857)
3	5.5	6.3929	(0.8929)
4	7.2	6.9000	0.3000
5	8	7.4071	0.5929
6	7.7	7.9143	(0.2143)
7	8.4	8.4214	(0.0214)

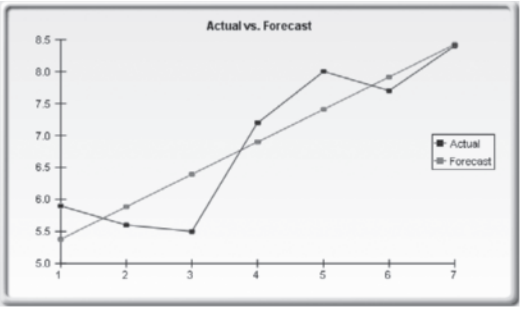


FIGURE 9.24 (Continued)

where the coefficient of determination is one less the ratio of the sums of squares of the errors (*SSE*) to the total sums of squares (*TSS*). In other words, the ratio of *SSE* to *TSS* is the unexplained portion of the analysis; thus, one less the ratio of *SSE* to *TSS* is the explained portion of the regression analysis.

Figure 9.25 provides a graphical explanation of the coefficient of determination. The estimated regression line is characterized by a series of predicted values ($\hat{Y}$); the average value of the dependent variable's data points is denoted $\overline{Y}$; and the individual data points are characterized by Y_i. Therefore, the total sum of squares, that is, the total variation in the data or the total variation about the average dependent value, is the total of the difference between the individual dependent values and its average (seen as the total squared distance of $Y_i - \overline{Y}$ in Figure 9.25). The explained sum of squares, the portion that is captured by the regression analysis, is the total of the difference between the regression's predicted value and the average dependent variable's data set (seen as the total squared distance of $\hat{Y} - \overline{Y}$ in Figure 9.25). The difference between the total variation (*TSS*) and the explained variation (*ESS*) is the unexplained sums of squares, also known as the sums of squares of the errors (*SSE*).

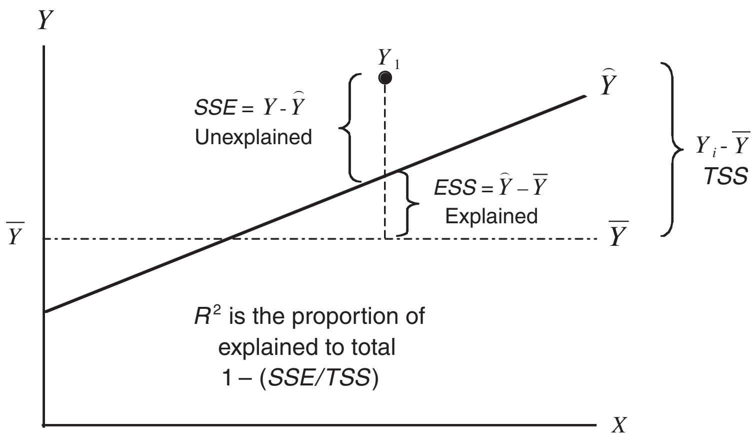


FIGURE 9.25 Explaining the coefficient of determination.

Another related statistic, the adjusted coefficient of determination, or the adjusted R-squared ($\bar{R}^2$), corrects for the number of independent variables (k) in a multivariate regression through a degrees-of-freedom correction to provide a more conservative estimate:

$$\bar{R}^2 = 1 - \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 \,/\, (k-2)}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2 \,/\, (k-1)} = 1 - \frac{SSE \,/\, (k-2)}{TSS \,/\, (k-1)}$$

The adjusted R-squared should be used instead of the regular R-squared in multivariate regressions because every time an independent variable is added into the regression analysis, the R-squared will increase, indicating that the percent variation explained has increased. This increase occurs even when nonsensical regressors are added. The adjusted R-squared takes the added regressors into account and penalizes the regression equation accordingly, providing a much better estimate of a regression model's goodness-of-fit.

Other goodness-of-fit statistics include the t-statistic and the F-statistic. The former is used to test if *each* of the estimated slope and intercept(s) is statistically significant, that is, if it is statistically significantly different from zero (therefore making sure that the intercept and slope estimates are statistically valid). The latter applies the same concepts but simultaneously for the entire regression equation including the intercept and slope(s). Using the previous example, the following illustrates how the t-statistic and F-statistic can be used in a regression analysis. (See the t-statistic and F-statistic tables at the end of the book for their corresponding critical values.) It is assumed that the reader is somewhat familiar with hypothesis testing and tests of significance in basic statistics.

Example Given the information from Excel's regression output in Figure 9.26, interpret the following:

1. Perform a hypothesis test on the slope and the intercept to see if they are *each* significant at a two-tailed alpha (α) of 0.05.

The null hypothesis H_o is such that the slope $\beta_1 = 0$ and the alternate hypothesis H_a is such that $\beta_1 \neq 0$. The t-statistic calculated is 4.6877, which exceeds the t-critical (2.9687 obtained from the t-statistic table at the end of this book) for a two-tailed alpha of 0.05 and degrees of freedom $n - k = 7 - 1 = 6$.[4] Therefore, the null hypothesis is rejected and one can state that the slope is statistically significantly different from 0, indicating that the regression's estimate of the slope is statistically significant. This hypothesis test can also be performed by looking at the t-statistic's corresponding p-value (0.0054), which is less than the alpha of 0.05, which means the null hypothesis is rejected.[5] The hypothesis test is then applied to the intercept, where the null hypothesis H_o is such that the intercept $\beta_0 = 0$ and the alternate hypothesis H_a is such that $\beta_0 \neq 0$. The t-statistic calculated is 7.4911, which exceeds the critical t value of 2.9687 for $n - k$ $(7 - 1 = 6)$ degrees of freedom, so, the null hypothesis is rejected, indicating that the intercept is statistically significantly different from 0, meaning that the regression's estimate of the intercept is statistically significant. The calculated p-value (0.0007) is also less than the alpha level, which means the null hypothesis is also rejected.

2. Perform a hypothesis test to see if both the slope and intercept are significant as a whole, in other words, if the estimated model is statistically significant at an alpha (α) of 0.05.

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	7.2014	7.2014	21.9747	0.0054
Residual	5	1.6386	0.3277		
Total	6	8.8400			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	4.3643	0.5826	7.4911	0.0007	2.8667	5.8619	2.8667	5.8619
X Variable 1	0.0845	0.0180	4.6877	0.0054	0.0382	0.1309	0.0382	0.1309

FIGURE 9.26 ANOVA and goodness-of-fit table.

Regression Statistics	
R-Squared (coefficient of determination)	0.8146
Adjusted R-Squared	0.7776
Multiple R (multiple correlation coefficient)	0.9026
Standard Error of the Estimates (SE_y)	0.5725
n Observations	7

Regression Results	Intercept	Ad Size
Coefficients	4.3643	0.0845
Standard Error	0.5826	0.0180
t-Statistic	7.4911	4.6877
p-Value	0.0007	*0.0054*
Lower 5%	2.8667	0.0382
Upper 95%	5.8619	0.1309

FIGURE 9.27 Additional regression output from Risk Simulator.

The simultaneous null hypothesis H_o is such that $\beta_0 = \beta_1 = 0$ and the alternate hypothesis H_a is $\beta_0 \neq \beta_1 \neq 0$. The calculated F-value is 21.9747, which exceeds the critical F-value (5.99 obtained from the table at the end of this book) for k (1) degrees of freedom in the numerator and n – k (7 – 1 = 6) degrees of freedom for the denominator, so the null hypothesis is rejected, indicating that both the slope and intercept are simultaneously significantly different from 0 and that the model as a whole is statistically significant. This result is confirmed by the p-value of 0.0054 (significance of F), which is less than the alpha value, thereby rejecting the null hypothesis and confirming that the regression as a whole is statistically significant.

3. Using Risk Simulator's regression output in Figure 9.27, interpret the R^2 value. How is it related to the correlation coefficient?

The calculated R^2 is 0.8146, meaning that 81.46 percent of the variation in the dependent variable can be explained by the variation in the independent variable. The R^2 is simply the square of the correlation coefficient, that is, the correlation coefficient between the independent and dependent variable is 0.9026.

Regression Assumptions

The following six assumptions are the requirements for a regression analysis to work:

1. The relationship between the dependent and independent variables is *linear*.

2. The expected value of the errors or *residuals is zero*.
3. The errors are *independently and normally distributed*.
4. The variance of the errors is constant or *homoskedastic* and not varying over time.
5. The errors are independent and *uncorrelated* with the explanatory variables.
6. The independent variables are uncorrelated to each other, meaning that no *multicollinearity* exists.

One very simple method to verify some of these assumptions is to use a scatter plot. This approach is simple to use in a bivariate regression scenario. If the assumption of the linear model is valid, the plot of the observed dependent variable values against the independent variable values should suggest a linear band across the graph with no obvious departures from linearity. Outliers may appear as anomalous points in the graph, often in the upper right-hand or lower left-hand corner of the graph. However, a point may be an outlier in either an independent or dependent variable without necessarily being far from the general trend of the data.

If the linear model is not correct, the shape of the general trend of the X–Y plot may suggest the appropriate function to fit (e.g., a polynomial, exponential, or logistic function). Alternatively, the plot may suggest a reasonable transformation to apply. For example, if the X–Y plot arcs from lower left to upper right so that data points either very low or very high in the independent variable lie below the straight line suggested by the data, while the middle data points of the independent variable lie on or above that straight line, taking square roots or logarithms of the independent variable values may promote linearity.

If the assumption of equal variances or homoskedasticity for the dependent variable is correct, the plot of the observed dependent variable values against the independent variable should suggest a band across the graph with roughly equal vertical width for all values of the independent variable. That is, the shape of the graph should suggest a tilted cigar and not a wedge or a megaphone.

A fan pattern like the profile of a megaphone, with a noticeable flare either to the right or to the left in the scatter plot, suggests that the variance in the values increases in the direction where the fan pattern widens (usually as the sample mean increases), and this in turn suggests that a transformation of the dependent variable values may be needed.

As an example, Figure 9.28 shows a scatter plot of two variables: sales revenue (dependent variable) and marketing costs (independent variable). Clearly, there is a positive relationship between the two variables, as is evident from the regression results in Figure 9.29, where the slope of the regression equation is a positive value (0.7447). The relationship is also

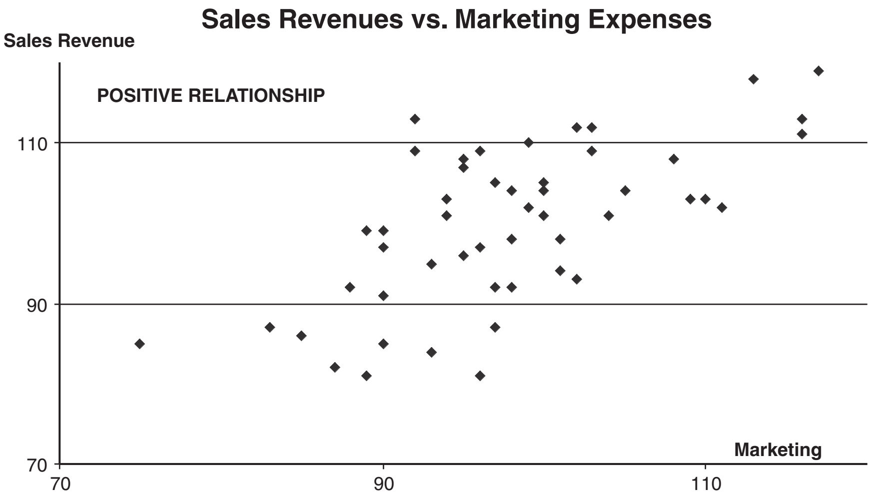


FIGURE 9.28 Scatter plot showing a positive relationship.

Summary:

Number of series: 2
Periods to forecast: 12
Seasonality: none
Error Measure: RMSE

Series: Sales Revenues

Method: Multiple Linear Regression

Statistics:

R-squared: 0.430
Adjusted R-squared: 0.4185
SSE: 2732.9
F Statistic: 36.263
F Probability: 2.32E-7
Durbin-Watson: 2.370
No. of Values: 50
Independent variables: 1 included out of 1 selected

Regression Variables:

Variable	Coefficient	t Statistic	Probability
Constant	26.897	2.215	0.03154
Marketing Expenses	0.7447	6.0219	2.32E-07

FIGURE 9.29 Bivariate regression results for positive relationship.

Summary:

Number of series: 3
Periods to forecast: 12
Seasonality: none
Error Measure: RMSE

HIGHER ADJUSTED R-SQUARED

Series: Sales Revenues

Method: Multiple Linear Regression

Statistics:
R-squared: 0.636
Adjusted R-squared: 0.6206
SSE: 1745.8
F Statistic: 41.079
F Probability: 4.81E-11
Durbin-Watson: 1.991
No. of Values: 50
Independent variables: 2 included out of 2 selected

SIGNIFICANT POSITIVE AND NEGATIVE RELATIONSHIPS

Regression Variables:

Variable	Coefficient	t Statistic	Probability
Constant	877.97	5.3085	2.94E-06
Marketing Expenses	0.6507	6.409	6.45E-08
Pricing	-8.1382	-5.155	4.97E-06

FIGURE 9.30 Multiple linear regression results for positive and negative relationships.

statistically significant at 0.05 alpha and the coefficient of determination is 0.43, indicating a somewhat weak but statistically significant relationship.

Compare that to a multiple linear regression in Figure 9.30, where another independent variable, pricing structure of the product, is added. The regression's adjusted coefficient of determination (adjusted R-squared) is now 0.62, indicating a much stronger regression model.[6] The pricing variable shows a negative relationship to the sales revenue, a very much expected result, as according to the law of demand in economics, a higher price point necessitates a lower quantity demanded, hence, lower sales revenues (this, of course, assumes an elastic demand curve). The t-statistics and corresponding probabilities (p-values) also indicate a statistically significant relationship.

In contrast, Figure 9.31 shows a scatter plot of two variables with little to no relationship, which is confirmed by the regression result in Figure 9.32, where the coefficient of determination is 0.066, close to being negligible. In addition, the calculated t-statistic and corresponding probability indicate that the marketing-expenses variable is statistically insignificant at the 0.05 alpha level, meaning that the regression equation is not significant (a fact that is also confirmed by the low F-statistic).

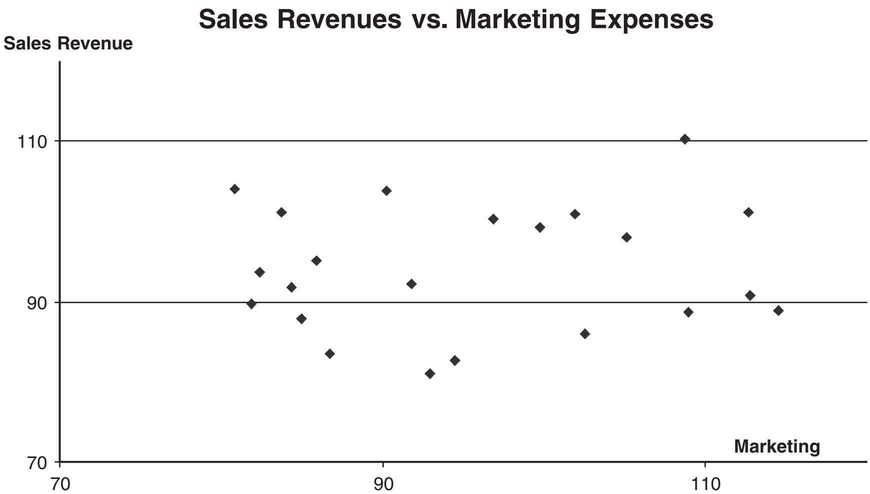


FIGURE 9.31 Scatter plot showing no relationship.

Summary:

Number of series: 2
Periods to forecast: 12
Seasonality: 12 months
Error Measure: RMSE

Series: Sales Revenues

Method: Multiple Linear Regression

Statistics:

R-squared: 0.066
Adjusted R-squared: 0.04622
SSE: 13661
F Statistic: 3.3743
F Probability: 0.07242
Durbin-Watson: 2.173
No. of Values: 50
Independent variables: 1 included out of 1 selected

LOW R-SQUARED IS AN INDICATION OF LITTLE TO NO RELATIONSHIP

Regression Variables:

Variable	Coefficient	t Statistic	Probability
Constant	82.966	6.0363	2.20E-07
Marketing Expenses	0.2265	1.8369	0.07242

FIGURE 9.32 Multiple regression results showing no relationship.

THE PITFALLS OF FORECASTING: OUTLIERS, NONLINEARITY, MULTICOLLINEARITY, HETEROSKEDASTICITY, AUTOCORRELATION, AND STRUCTURAL BREAKS

Other than being good modeling practice to create scatter plots prior to performing regression analysis, the scatter plot can also sometimes, on a fundamental basis, provide significant amounts of information regarding the behavior of the data series. Blatant violations of the regression assumptions can be spotted easily and effortlessly, without the need for more detailed and fancy econometric specification tests. For instance, Figure 9.33 shows the existence of outliers. Figure 9.34's regression results, which include the outliers, indicate that the coefficient of determination is only 0.252 as compared to 0.447 in Figure 9.35 when the outliers are removed.

Values may not be identically distributed because of the presence of outliers. Outliers are anomalous values in the data. Outliers may have a strong influence over the fitted slope and intercept, giving a poor fit to the bulk of the data points. Outliers tend to increase the estimate of residual variance, lowering the chance of rejecting the null hypothesis. They may be due to recording errors, which may be correctable, or they may be due to the dependent-variable values not all being sampled from the same population. Apparent outliers may also be due to the dependent-variable values being from the same, but nonnormal, population. Outliers may show up clearly in an X–Y scatter plot of the data, as points that do not lie near the general linear

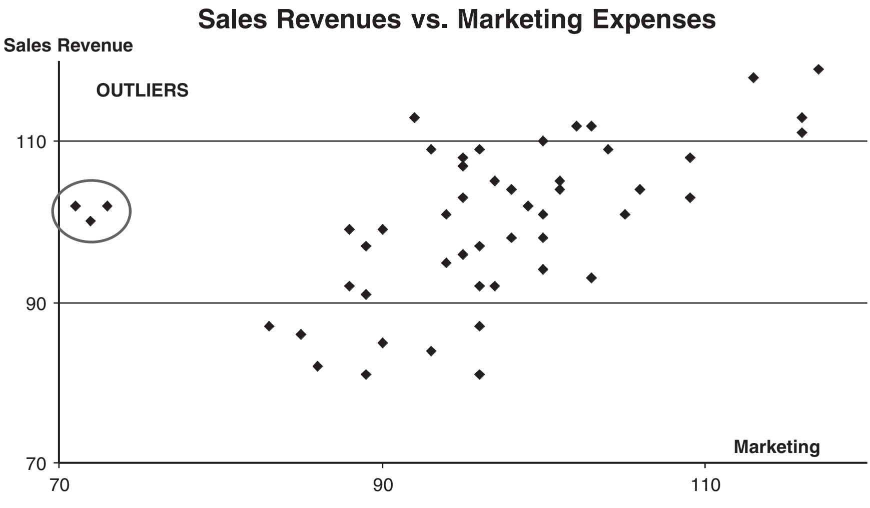


FIGURE 9.33 Scatter plot showing outliers.

Summary:

Number of series: 2
Periods to forecast: 12
Seasonality: 12 months
Error Measure: RMSE

Series: Sales Revenues

Method: Multiple Linear Regression

Statistics:
R-squared: 0.252
Adjusted R-squared: 0.2367
SSE: 3417.6
F Statistic: 16.198
F Probability: 2.01E-4
Durbin-Watson: 1.945
No. of Values: 50
Independent variables: 1 included out of 1 selected

Regression Variables:

Variable	Coefficient	t Statistic	Probability
Constant	53.269	4.5619	3.51E-05
Marketing Expenses	0.4857	4.0247	2.01E-04

FIGURE 9.34 Regression results with outliers.

trend of the data. A point may be an unusual value in either an independent or dependent variable without necessarily being an outlier in the scatter plot.

The method of least squares involves minimizing the sum of the squared vertical distances between each data point and the fitted line. Because of this, the fitted line can be highly sensitive to outliers. In other words, least squares regression is not resistant to outliers; thus, neither is the fitted-slope estimate. A point vertically removed from the other points can cause the fitted line to pass close to it, instead of following the general linear trend of the rest of the data, especially if the point is relatively far horizontally from the center of the data (the point represented by the mean of the independent variable and the mean of the dependent variable). Such points are said to have high leverage: the center acts as a fulcrum, and the fitted line pivots toward high-leverage points, perhaps fitting the main body of the data poorly. A data point that is extreme in dependent variables but lies near the center of the data horizontally will not have much effect on the fitted slope, but by changing the estimate of the mean of the dependent variable, it may affect the fitted estimate of the intercept.

However, great care should be taken when deciding if the outliers should be removed. Although in most cases when outliers are removed, the regression

Summary:

Number of series: 2
Periods to forecast: 12
Seasonality: 12 months
Error Measure: RMSE

Series: Sales Revenues

Method: Multiple Linear Regression

Statistics:

R-squared: 0.447
Adjusted R-squared: 0.4343
SSE: 2524.9
F Statistic: 36.321
F Probability: 2.84E-7
Durbin-Watson: 2.242
No. of Values: 47
Independent variables: 1 included out of 1 selected

COMPARE R-SQUARED BETWEEN REGRESSION WITH OUTLIERS AND WITHOUT OUTLIERS!

Regression Variables:

Variable	Coefficient	t Statistic	Probability
Constant	19.447	1.4512	0.1537
Marketing Expenses	0.8229	6.0267	2.84E-07

FIGURE 9.35 Regression results with outliers deleted.

results look better, a priori justification must first exist. For instance, if one is regressing the performance of a particular firm's stock returns, outliers caused by downturns in the stock market should be included; these are not truly outliers as they are inevitabilities in the business cycle. Forgoing these outliers and using the regression equation to forecast one's retirement fund based on the firm's stocks will yield incorrect results at best. In contrast, suppose the outliers are caused by a single nonrecurring business condition (e.g., merger and acquisition) and such business structural changes are not forecast to recur, then these outliers should be removed and the data cleansed prior to running a regression analysis.

Figure 9.36 shows a scatter plot with a nonlinear relationship between the dependent and independent variables. In a situation such as the one in Figure 9.36, a linear regression will not be optimal. A nonlinear transformation should first be applied to the data before running a regression. One simple approach is to take the natural logarithm of the independent variable (other approaches include taking the square root or raising the independent variable to the second or third power) and regress the sales revenue on this transformed marketing-cost data series. Figure 9.37 shows the regression results with a coefficient of determination at 0.938, as compared to 0.707 in

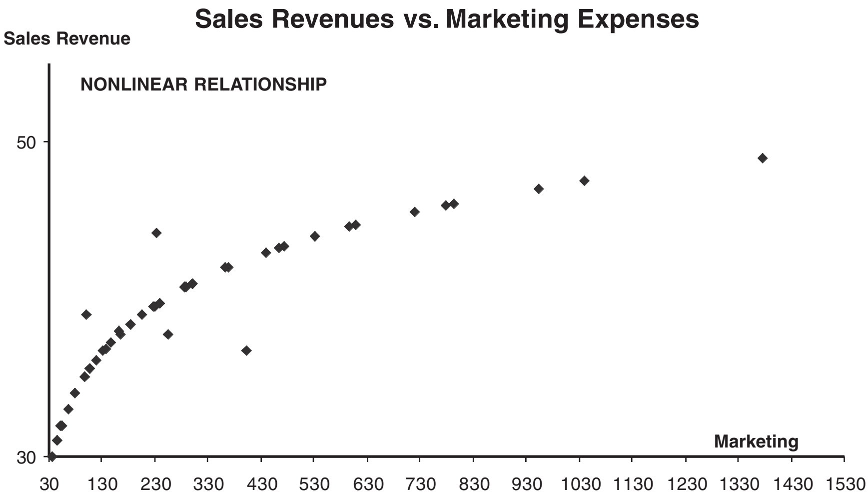


FIGURE 9.36 Scatter plot showing a nonlinear relationship.

Summary:

Number of series: 3
Periods to forecast: 12
Seasonality: none
Error Measure: RMSE

Series: Sales Revenues

Method: Multiple Linear Regression

Statistics:

R-squared: 0.938
Adjusted R-squared: 0.9364
SSE: 101.74
F Statistic: 722.25
F Probability: 1.39E-30
Durbin-Watson: 1.825
No. of Values: 50
Independent variables: 1 included out of 1 selected

Regression Variables:

Variable	Coefficient	t Statistic	Probability
Constant	10.208	9.6141	9.03E-13
Nonlinear Marketing Expenses	5.3783	26.875	1.39E-30

FIGURE 9.37 Regression results using a nonlinear transformation.

Summary:

Number of series: 3
Periods to forecast: 12
Seasonality: none
Error Measure: RMSE

Series: Sales Revenues

Method: Multiple Linear Regression

Statistics:

R-squared: 0.707
Adjusted R-squared: 0.7013
SSE: 477.72
F Statistic: 116.04
F Probability: 2.09E-14
Durbin-Watson: 0.992
No. of Values: 50
Independent variables: 1 included out of 1 selected

Regression Variables:

Variable	Coefficient	t Statistic	Probability
Constant	33.358	52.658	4.00E-44
Linear Marketing Expenses	0.01639	10.772	2.09E-14

FIGURE 9.38 Regression results using linear data.

Figure 9.38 when a simple linear regression is applied to the original data series without the nonlinear transformation.

If the linear model is not the correct one for the data, then the slope and intercept estimates and the fitted values from the linear regression will be biased, and the fitted slope and intercept estimates will not be meaningful. Over a restricted range of independent or dependent variables, nonlinear models may be well approximated by linear models (this is in fact the basis of linear interpolation), but for accurate prediction a model appropriate to the data should be selected. An examination of the X–Y scatter plot may reveal whether the linear model is appropriate. If there is a great deal of variation in the dependent variable, it may be difficult to decide what the appropriate model is. In this case, the linear model may do as well as any other, and has the virtue of simplicity. Refer to Appendix C—Detecting and Fixing Heteroskedasticity—for specification tests of nonlinearity and heteroskedasticity as well as ways to fix them.

However, great care should be taken here as both the original linear data series of marketing costs should not be added with the nonlinearly transformed marketing costs in the regression analysis. Otherwise, multicollinearity occurs; that is, marketing costs are highly correlated to the natural logarithm of marketing costs, and if both are used as independent variables in a multivariate regression analysis, the assumption of no multicollinearity is violated and the regression analysis breaks down. Figure 9.39 illustrates what happens when multicollinearity strikes. Notice that the coefficient of determination (0.938) is the same as the nonlinear transformed regression (Figure 9.37). However, the adjusted coefficient of determination went down from 0.9364 (Figure 9.37) to 0.9358 (Figure 9.39). In addition, the previously statistically significant marketing-costs variable in Figure 9.38 now becomes insignificant (Figure 9.39) with a probability value increasing from close to zero to 0.4661. A basic symptom of multicollinearity is low t-statistics coupled with a high R-squared (Figure 9.39). See Appendix D—Detecting and Fixing Multicollinearity—for further details on detecting multicollinearity in a regression.

Another common violation is heteroskedasticity, that is, the variance of the errors increases over time. Figure 9.40 illustrates this case, where the width of the vertical data fluctuations increases or fans out over time. In this example, the data points have been changed to exaggerate the effect. However, in most time-series analysis, checking for heteroskedasticity is a much

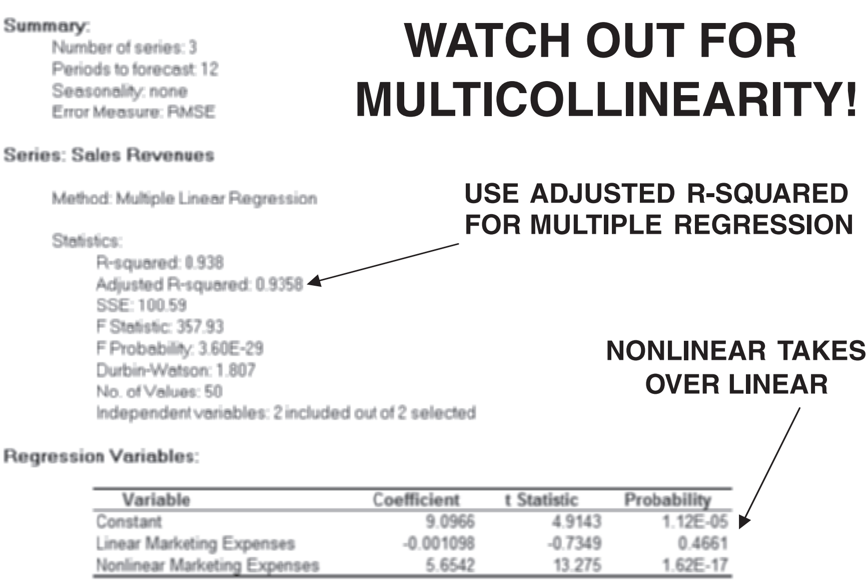


FIGURE 9.39 Regression results using both linear and nonlinear transformations.

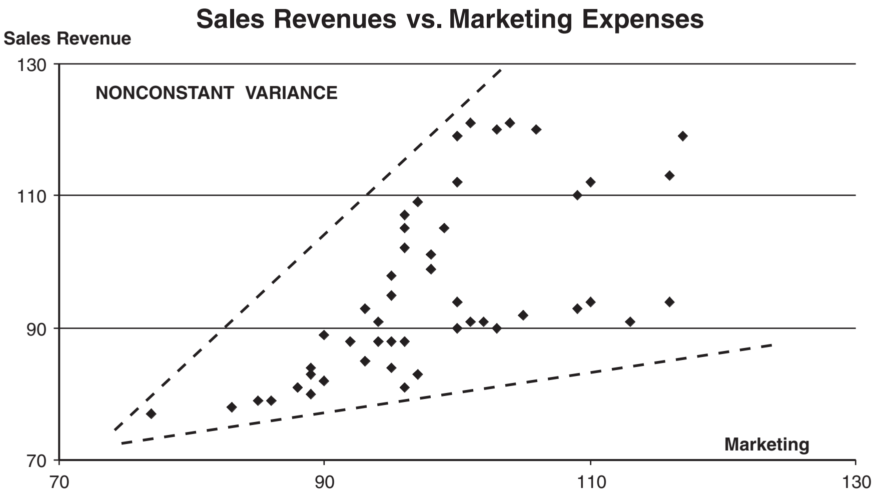


FIGURE 9.40 Scatter plot showing heteroskedasticity with nonconstant variance.

more difficult task. See Appendix C—Detecting and Fixing Heteroskedasticity—for further details. And correcting for heteroskedasticity is an even greater challenge.[7] Notice in Figure 9.41 that the coefficient of determination drops significantly when heteroskedasticity exists. As is, the current regression model is insufficient and incomplete. Refer to Appendix C for more details.

If the variance of the dependent variable is not constant, then the error's variance will not be constant. The most common form of such heteroskedasticity in the dependent variable is that the variance of the dependent variable may increase as the mean of the dependent variable increases for data with positive independent and dependent variables.

Unless the heteroskedasticity of the dependent variable is pronounced, its effect will not be severe: The least-squares estimates will still be unbiased, and the estimates of the slope and intercept will either be normally distributed if the errors are normally distributed, or at least normally distributed asymptotically (as the number of data points becomes large) if the errors are not normally distributed. The estimate for the variance of the slope and overall variance will be inaccurate, but the inaccuracy is not likely to be substantial if the independent-variable values are symmetric about their mean.

Heteroskedasticity of the dependent variable is usually detected informally by examining the X–Y scatter plot of the data before performing the regression. If both nonlinearity and unequal variances are present, employing a transformation of the dependent variable may have the effect of simultaneously improving the linearity and promoting equality of the variances.

Summary:

Number of series: 2
Periods to forecast: 12
Seasonality: 12 months
Error Measure: RMSE

WATCH OUT FOR HETEROSKEDASTICITY!

Series: Sales Revenues

Method: Multiple Linear Regression

Statistics:
R-squared: 0.398
Adjusted R-squared: 0.3858
SSE: 5190.1
F Statistic: 31.777
F Probability: 8.94E-7
Durbin-Watson: 2.755
No. of Values: 50
Independent variables: 1 included out of 1 selected

Regression Variables:

Variable	Coefficient	t Statistic	Probability
Constant	1.5742	0.09421	0.9253
Marketing Expenses	0.9586	5.6371	8.94E-07

FIGURE 9.41 Regression results with heteroskedasticity.

Otherwise, a weighted least-squares linear regression may be the preferred method of dealing with nonconstant variance of the dependent variable.

OTHER TECHNICAL ISSUES IN REGRESSION ANALYSIS

If the data to be analyzed by linear regression violate one or more of the linear regression assumptions, the results of the analysis may be incorrect or misleading. For example, if the assumption of independence is violated, then linear regression is not appropriate. If the assumption of normality is violated or outliers are present, then the linear regression goodness-of-fit test may not be the most powerful or informative test available, and this could mean the difference between detecting a linear fit or not. A nonparametric, robust, or resistant regression method, a transformation, a weighted least-squares linear regression, or a nonlinear model may result in a better fit. If the population variance for the dependent variable is not constant, a weighted least-squares linear regression or a transformation of the dependent variable may provide a means of fitting a regression adjusted for the inequality of the variances. Often, the impact of an assumption violation on the linear regression result depends on the extent of the violation (such as how nonconstant the variance of the dependent variable is, or how skewed the dependent variable population distribution is). Some small violations may have little practical effect on the analysis, while other violations may render the linear regression result useless and incorrect. Other potential assumption violations include:

- Lack of independence in the dependent variable.
- Independent variable is random, not fixed.
- Special problems with few data points.
- Special problems with regression through the origin.

Lack of Independence in the Dependent Variable

Whether the independent-variable values are independent of each other is generally determined by the structure of the experiment from which they arise. The dependent-variable values collected over time may be autocorrelated. For serially correlated dependent-variable values, the estimates of the slope and intercept will be unbiased, but the estimates of their variances will not be reliable and hence the validity of certain statistical goodness-of-fit tests will be flawed. An ARIMA model may be better in such circumstances.

The Independent Variable Is Random, Not Fixed

The usual linear regression model assumes that the observed independent variables are fixed, not random. If the independent values are not under the control of the experimenter (i.e., are observed but not set), and if there is in fact underlying variance in the independent variable, but they have the same variance, the linear model is called an errors-in-variables model or structural model. The least-squares fit will still give the best linear predictor of the dependent variable, but the estimates of the slope and intercept will be biased (will not have expected values equal to the true slope and variance). A stochastic forecast model may be a better alternative here.

Special Problems with Few Data Points (Micronumerosity)

If the number of data points is small (also termed *micronumerosity*), it may be difficult to detect assumption violations. With small samples, assumption violations such as nonnormality or heteroskedasticity of variances are difficult to detect even when they are present. With a small number of data points, linear regression offers less protection against violation of assumptions. With few data points, it may be hard to determine how well the fitted line matches the data, or whether a nonlinear function would be more appropriate.

Even if none of the test assumptions are violated, a linear regression on a small number of data points may not have sufficient power to detect a significant difference between the slope and zero, even if the slope is nonzero. The power depends on the residual error, the observed variation in the independent variable, the selected significance alpha level of the test, and the number of data points. Power decreases as the residual variance increases, decreases as the significance level is decreased (i.e., as the test is made more stringent), increases as the variation in the observed independent variable increases, and increases as the number of data points increases. If a statistical significance test with a small number of data points produces a surprisingly nonsignificant probability value, then lack of power may be the reason. The best time to avoid such problems is in the design stage of an experiment, when appropriate minimum sample sizes can be determined, perhaps in consultation with an econometrician, before data collection begins.

Special Problems with Regression Through the Origin

The effects of nonconstant variance of the dependent variable can be particularly severe for a linear regression when the line is forced through the origin: The estimate of variance for the fitted slope may be much smaller than the actual variance, making the test for the slope nonconservative (more likely to reject the null hypothesis that the slope is zero than what the stated significance level indicates). In general, unless there is a structural or theoretical reason to assume that the intercept is zero, it is preferable to fit both the slope and intercept.

APPENDIX A—FORECAST INTERVALS

The forecast interval estimated in a forecast (an approach also used by Risk Simulator) is illustrated in Figure 9.42. The confidence interval (CI) is estimated by

$$\hat{Y}_i \pm Z\left[\frac{RMSE}{N-T}\right]N$$

where $\hat{Y}_i$ is the ith forecast estimate; Z is the standard-normal statistic (see the standard-normal tables at the end of this book); $RMSE$ is the root mean-squared error previously calculated; N is the number of historical data points; and T is the forecast period. When N is a relatively small number

Period	Raw Data
1	265.22
2	146.64
3	182.50
4	118.54
5	180.04
6	167.45
7	231.75
8	223.71
9	192.98
10	122.29
11	336.65
12	186.50
13	194.27
14	149.19
15	210.06
36	647.82
37	650.81
38	677.54
39	666.56

Forecast Values

Forecast	5%	95%
710.07	586.91	833.23
701.52	575.03	828.01
756.04	626.04	886.04
818.99	685.27	952.71
794.37	656.71	932.02
Estimated RMSE		72.951

$$CI = \hat{Y} \pm Z \left[\frac{RMSE}{(N-T)} N \right]$$

Period	Forecast (T)	Stdev	Z-statistic	Lower	Upper		
40	1	74.87	1.645	586.91	833.23	***RMSE***	72.951
41	2	76.89	1.645	575.03	828.01	***Data Points (N)***	39
42	3	79.03	1.645	626.03	886.04		
43	4	81.29	1.645	685.27	952.71		
44	5	83.68	1.645	656.71	932.02		

FIGURE 9.42 Forecast interval estimation.

(usually less than 30), then the same analysis can be performed using the t-statistic in place of the Z-value (see the t-statistic table at the end of this book).

Clearly, this approach is a modification of the more common confidence interval estimate of

$$\hat{Y}_i \pm Z \frac{\sigma}{\sqrt{n}}$$

applicable within a data set. Here, it is assumed that

$$\left[\frac{RMSE}{N-T} \right] N \approx \frac{\sigma}{\sqrt{n}}$$

and the inclusion of the T variable is simply to adjust for the added degrees of freedom when forecasting outside of the original data set.

APPENDIX B—ORDINARY LEAST SQUARES

The following illustrates the concept of the ordinary least-squares regression line. Figure 9.43 shows the data on the dependent variable (Y) and independent variable (X) as well as the results estimated using Excel's solver add-in.

	A	B	C	D	E	F	G	H
1	Y	X	Slope	Intercept	Predicted	Residual	Squared Resid	
2	1000	3	91.98	2489.16	2765.09	1765.09	3115530.48	
3	3333	3	91.98	2489.16	2765.09	-567.91	322525.70	
4	2222	3	91.98	2489.16	2765.09	543.09	294942.99	
5	1111	2	91.98	2489.16	2673.11	1562.11	2440188.73	
6	5555	3	91.98	2489.16	2765.09	-2789.91	7783617.14	
7	2222	2	91.98	2489.16	2673.11	451.11	203500.54	
8	2222	3	91.98	2489.16	2765.09	543.09	294942.99	
9	5555	3	91.98	2489.16	2765.09	-2789.91	7783617.14	
10	4444	7	91.98	2489.16	3132.99	-1311.01	1718743.79	
11	3333	6	91.98	2489.16	3041.02	-291.98	85255.17	
12	2222	7	91.98	2489.16	3132.99	910.99	829905.16	
13	1111	8	91.98	2489.16	3224.97	2113.97	4468858.59	
14	5555	7	91.98	2489.16	3132.99	-2422.01	5866126.11	
15	2222	6	91.98	2489.16	3041.02	819.02	670785.76	
16	2222	7	91.98	2489.16	3132.99	910.99	829905.16	
17	5555	6	91.98	2489.16	3041.02	-2513.98	6320120.01	
18	4444	5	91.98	2489.16	2949.04	-1494.96	2234908.63	
19	1111	6	91.98	2489.16	3041.02	1930.02	3724958.34	
20	2222	4	91.98	2489.16	2857.06	635.06	403304.67	
21	3333	5	91.98	2489.16	2949.04	-383.96	147426.11	
22	2222	4	91.98	2489.16	2857.06	635.06	403304.67	
23	1111	4	91.98	2489.16	2857.06	1746.06	3048735.05	

Optimization Parameter	
Intercept	2489.16
Slope	91.98
Sum of Squared Residuals	52991202.91

Excel Estimated Parameter	
Intercept	2489.16
Slope	91.98

FIGURE 9.43 Using optimization to estimate regression intercept and slope.

Arbitrary starting points of the slope and intercept values are fitted back into the data points and the squared residuals are calculated. Then, the optimal slope and intercept values are calculated through minimizing the sum of the squared residuals.

To get started, make sure Excel's Solver is added in by clicking on *Tools | Add-Ins*. Verify that the check-box beside *Solver Add-In* is selected (Figure 9.44). Then, back in the Excel model, click on *Tools | Solver* and make sure the *sum of squared residuals* (cell E28) is set as the target cell to minimize through systematically changing the intercept and slope values (cells E26 and E27) as seen in Figure 9.45.

Solving yields an intercept value of 2489.16 and a slope of 91.98. These results can be verified using Excel's built-in *slope* and *intercept* functions (Figure 9.46). In other words, the ordinary least-squares regression equation approach is the unique line (as described by an intercept and slope) that minimizes all possible vertical errors (total sum of squared residuals), making it the best-fitting line through a data set.

FIGURE 9.44 Excel Solver add-in.

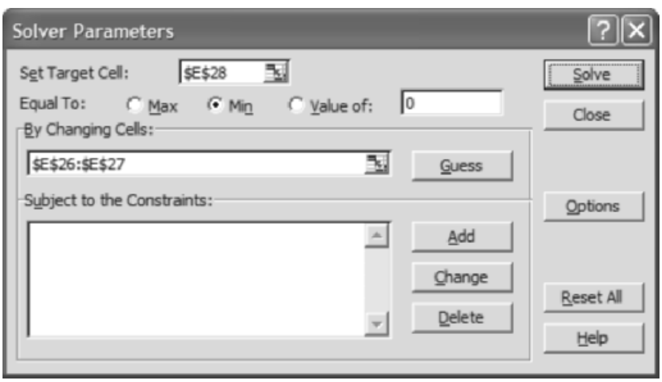


FIGURE 9.45 Excel Solver parameters.

Optimization Parameters	
Intercept	2489.16
Slope	91.98
Sum of Squared Residuals	52991202.91

Excel Estimated Parameter	
Slope	2489.16
Intercept	91.98

FIGURE 9.46 Optimized ordinary least squares results.

APPENDIX C—DETECTING AND FIXING HETEROSKEDASTICITY

Several tests exist to test for the presence of heteroskedasticity. These tests also are applicable for testing misspecifications and nonlinearities. The simplest approach is to graphically represent each independent variable against the dependent variable as illustrated earlier in the chapter. Another approach is to apply one of the most widely used models, the White's test, where the test is based on the null hypothesis of no heteroskedasticity against an alternate hypothesis of heteroskedasticity of some unknown general form. The test statistic is computed by an auxiliary or secondary regression, where the squared residuals or errors from the first regression are regressed on all possible (and nonredundant) cross products of the regressors. For example, suppose the following regression is estimated:

$$Y = \beta_0 + \beta_1 X + \beta_2 Z + \varepsilon_t$$

The test statistic is then based on the auxiliary regression of the errors (ε):

$$\varepsilon_t^2 = \alpha_0 + \alpha_1 X + \alpha_2 Z + \alpha_3 X^2 + \alpha_4 Z^2 + \alpha_5 XZ + v_t$$

The nR^2 statistic is the White's test statistic, computed as the number of observations (n) times the centered R-squared from the test regression. White's test statistic is asymptotically distributed as a χ^2 with degrees of freedom equal to the number of independent variables (excluding the constant) in the test regression.

The White's test is also a general test for model misspecification, because the null hypothesis underlying the test assumes that the errors are both homoskedastic and independent of the regressors, and that the linear specification of the model is correct. Failure of any one of these conditions could lead to a significant test statistic. Conversely, a nonsignificant test statistic implies that none of the three conditions is violated. For instance, the resulting F-statistic is an omitted variable test for the joint significance of all cross products, excluding the constant.

One method to fix heteroskedasticity is to make it homoskedastic by using a weighted least-squares (WLS) approach. For instance, suppose the following is the original regression equation:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$$

Further suppose that X_2 is heteroskedastic. Then transform the data used in the regression into:

$$Y = \frac{\beta_0}{X_2} + \beta_1 \frac{X_1}{X_2} + \beta_2 + \beta_3 \frac{X_3}{X_2} + \frac{\varepsilon}{X_2}$$

The model can be redefined as the following WLS regression:

$$Y_{WLS} = \beta_0{}^{WLS} + \beta_1{}^{WLS} X_1 + \beta_2{}^{WLS} X_2 + \beta_3{}^{WLS} X_3 + v$$

Alternatively, the Park's test can be applied to test for heteroskedasticity and to fix it. The Park's test model is based on the original regression equation, uses its errors, and creates an auxiliary regression that takes the form of:

$$\ln e_i^2 = \beta_1 + \beta_2 \ln X_{k,i}$$

Suppose β_2 is found to be statistically significant based on a t-test, then heteroskedasticity is found to be present in the variable $X_{k,i}$. The remedy therefore is to use the following regression specification:

$$\frac{Y}{\sqrt{X_k^{\beta_2}}} = \frac{\beta_1}{\sqrt{X_k^{\beta_2}}} + \frac{\beta_2 X_2}{\sqrt{X_k^{\beta_2}}} + \frac{\beta_3 X_3}{\sqrt{X_k^{\beta_2}}} + \varepsilon$$

APPENDIX D—DETECTING AND FIXING MULTICOLLINEARITY

Multicollinearity exists when there is a linear relationship between the independent variables. When this occurs, the regression equation cannot be estimated at all. In near collinearity situations, the estimated regression equation will be biased and provide inaccurate results. This situation is especially true when a step-wise regression approach is used, where the statistically significant independent variables will be thrown out of the regression mix earlier than expected, resulting in a regression equation that is neither efficient nor accurate.

As an example, suppose the following multiple regression analysis exists, where

$$Y_i = \beta_1 + \beta_2 X_{2,i} + \beta_3 X_{3,i} + \varepsilon_i$$

then the estimated slopes can be calculated through

$$\hat{\beta}_2 = \frac{\sum Y_i X_{2,i} \sum X_{3,i}^2 - \sum Y_i X_{3,i} \sum X_{2,i} X_{3,i}}{\sum X_{2,i}^2 \sum X_{3,i}^2 - \left(\sum X_{2,i} X_{3,i}\right)^2}$$

$$\hat{\beta}_3 = \frac{\sum Y_i X_{3,i} \sum X_{2,i}^2 - \sum Y_i X_{2,i} \sum X_{2,i} X_{3,i}}{\sum X_{2,i}^2 \sum X_{3,i}^2 - \left(\sum X_{2,i} X_{3,i}\right)^2}$$

Now suppose that there is perfect multicollinearity, that is, there exists a perfect linear relationship between X_2 and X_3, such that $X_{3,i} = \lambda X_{2,i}$ for all positive values of λ. Substituting this linear relationship into the slope calculations for β_2, the result is indeterminate. In other words, we have

$$\hat{\beta}_2 = \frac{\sum Y_i X_{2,i} \sum \lambda^2 X_{2,i}^2 - \sum Y_i \lambda X_{2,i} \sum \lambda X_{2,i}^2}{\sum X_{2,i}^2 \sum \lambda^2 X_{2,i}^2 - \left(\sum \lambda \ X_{2,i}^2\right)^2} = \frac{0}{0}$$

The same calculation and results apply to β_3, which means that the multiple regression analysis breaks down and cannot be estimated given a perfect collinearity condition.

One quick test of the presence of multicollinearity in a multiple regression equation is that the R-squared value is relatively high while the t-statistics are relatively low. (See Figure 9.39 for an illustration of this effect.) Another quick test is to create a correlation matrix between the independent variables. A high cross correlation indicates a potential for multicollinearity. The rule of thumb is that a correlation with an absolute value greater than 0.75 is indicative of severe multicollinearity.

Another test for multicollinearity is the use of the variance inflation factor (VIF), obtained by regressing each independent variable to all the other independent variables, obtaining the R-squared value and calculating the VIF of that variable by estimating:

$$VIF_i = \frac{1}{(1 - R_i^2)}$$

A high VIF value indicates a high R-squared near unity. As a rule of thumb, a VIF value greater than 10 is usually indicative of destructive multicollinearity.

APPENDIX E—DETECTING AND FIXING AUTOCORRELATION

One very simple approach to test for autocorrelation is to graph the time series of a regression equation's residuals. If these residuals exhibit some cyclicality, then autocorrelation exists. Another more robust approach to detect autocorrelation is the use of the Durbin–Watson statistic, which estimates the potential for a first-order autocorrelation. The Durbin–Watson test also identifies model misspecification, that is, if a particular time-series variable is correlated to itself one period prior. Many time-series data tend to be autocorrelated to their historical occurrences. This relationship can be due to multiple reasons, including the variables' spatial relationships (similar time and space), prolonged economic shocks and events, psychological inertia, smoothing, seasonal adjustments of the data, and so forth.

The Durbin–Watson statistic is estimated by the sum of the squares of the regression errors for one period prior to the sum of the current period's errors:

$$DW = \frac{\sum(\varepsilon_t - \varepsilon_{t-1})^2}{\sum \varepsilon_t^2}$$

There is a Durbin–Watson critical statistic table at the end of the book that provides a guide as to whether a statistic implies any autocorrelation.

Another test for autocorrelation is the Breusch–Godfrey test, where for a regression function in the form of:

$$Y = f(X_1, X_2, \ldots, X_k)$$

estimate this regression equation and obtain its errors ε_t. Then, run the secondary regression function in the form of:

$$Y = f(X_1, X_2, \ldots, X_k\ \varepsilon_{t-1}, \varepsilon_{t-2}, \varepsilon_{t-p})$$

Obtain the R-squared value and test it against a null hypothesis of no autocorrelation versus an alternate hypothesis of autocorrelation, where the test statistic follows a chi-square distribution of p degrees of freedom:

$$R^2(n-p) \sim \chi^2_{df=p}$$

Fixing autocorrelation requires more advanced econometric models including the applications of ARIMA (Autoregressive Integrated Moving Average) or ECM (Error Correction Models). However, one simple fix is to take the lags of the dependent variable for the appropriate periods, add them into the regression function, and test for their significance, for instance:

$$Y_t = f(Y_{t-1}, Y_{t-2}, \ldots, Y_{t-p}, X_1, X_2, \ldots, X_k)$$

QUESTIONS

1. Explain what each of the following terms means:
   a. Time-series analysis
   b. Ordinary least squares
   c. Regression analysis
   d. Heteroskedasticity
   e. Autocorrelation
   f. Multicollinearity
   g. ARIMA
2. What is the difference between the R-squared versus the adjusted R-squared measure in a regression analysis? When is each applicable and why?
3. Explain why if each of the following is not detected properly or corrected for in the model, the estimated regression model will be flawed:
   a. Heteroskedasticity
   b. Autocorrelation
   c. Multicollinearity
4. Explain briefly how to fix the problem of nonlinearity in the data set.

EXERCISE

1. Based on the data in the chapter examples, re-create the following using Excel:
   a. Double-moving average model
   b. Single exponential-smoothing model
   c. Additive seasonality model
   d. Holt–Winters multiplicative model

PART Six

Risk Diversification

CHAPTER 10

The Search for the Optimal Decision

In most simulation models, there are variables over which you have control, such as how much to charge for a product or how much to invest in a project. These controlled variables are called decision variables. Finding the optimal values for decision variables can make the difference between reaching an important goal and missing that goal. This chapter details the optimization process at a high-level, while Chapter 11, Optimization under Uncertainty, provides two step-by-step examples of resource optimization and portfolio optimization solved using the Risk Simulator software.

WHAT IS AN OPTIMIZATION MODEL?

In today's highly competitive global economy, companies are faced with many difficult decisions. These decisions include allocating financial resources, building or expanding facilities, managing inventories, and determining product-mix strategies. Such decisions might involve thousands or millions of potential alternatives. Considering and evaluating each of them would be impractical or even impossible. A model can provide valuable assistance in incorporating relevant variables when analyzing decisions and finding the best solutions for making decisions. Models capture the most important features of a problem and present them in a form that is easy to interpret. Models often provide insights that intuition alone cannot. An optimization model has three major elements: decision variables, constraints, and an objective. In short, the optimization methodology finds the best combination or permutation of decision variables (e.g., which products to sell and which projects to execute) in every conceivable way such that the objective is maximized (e.g., revenues and net income) or minimized (e.g., risk and costs) while still satisfying the constraints (e.g., budget and resources).

Obtaining optimal values generally requires that you search in an iterative or ad hoc fashion. This search involves running one iteration for an initial set of values, analyzing the results, changing one or more values, rerunning the model, and repeating the process until you find a satisfactory solution.

This process can be very tedious and time consuming even for small models, and often it is not clear how to adjust the values from one iteration to the next.

A more rigorous method systematically enumerates all possible alternatives. This approach guarantees optimal solutions if the model is correctly specified. Suppose that an optimization model depends on only two decision variables. If each variable has 10 possible values, trying each combination requires 100 iterations (10^2 alternatives). If each iteration is very short (e.g., 2 seconds), then the entire process could be done in approximately three minutes of computer time.

However, instead of two decision variables, consider six, then consider that trying all combinations requires 1,000,000 iterations (10^6 alternatives). It is easily possible for complete enumeration to take weeks, months, or even years to carry out (see Figure 10.1).

THE TRAVELING FINANCIAL PLANNER

A very simple example is in order. Figure 10.2 illustrates the traveling financial planner problem. Suppose the traveling financial planner has to make three sales trips to New York, Chicago, and Seattle. Further suppose that the order of arrival at each city is irrelevant. All that is important in this simple example is to find the lowest total cost possible to cover all three cities. Figure 10.2 also lists the flight costs from these different cities.

The problem here is cost minimization, suitable for optimization. One basic approach to solving this problem is through an ad hoc or brute force

An approach used to find the combination of inputs to achieve the best possible output subject to satisfying certain prespecified conditions

- What stocks to pick in a portfolio, as well as the weights of each stock as a percentage of total budget
- Optimal staffing needs for a production line
- Project and strategy selection and prioritization
- Inventory optimization
- Optimal pricing and royalty rates
- Utilization of employees for workforce planning
- Configuration of machines for production scheduling
- Location of facilities for distribution
- Tolerances in manufacturing design
- Treatment policies in waste management

FIGURE 10.1 What is optimization?

- You have to travel and visit clients in New York, Chicago, and Seattle
- You may start from any city and you will stay at your final city, that is, you will need to purchase three airline tickets
- Your goal is to travel as cheaply as possible given these rates:

Route	Airfare
Seattle – Chicago	\$325
Chicago – Seattle	\$225
New York – Seattle	\$350
Seattle – New York	\$375
Chicago – New York	\$325
New York – Chicago	\$325

- How do you solve the problem?
  - Ad Hoc approach - start trying different combinations
  - Enumeration - look at all possible alternatives

FIGURE 10.2 Traveling financial planner problem.

method, that is, manually list all six possible permutations as seen in Figure 10.3. Clearly the cheapest itinerary is going from the East Coast to the West Coast, going from New York to Chicago and finally on to Seattle.[1] Here, the problem is simple and can be calculated manually, as there are three cities and hence six possible itineraries.[2] However, add two more cities and the total number of possible itineraries jumps to 120.[3] Performing an ad hoc calculation will be fairly intimidating and time consuming. On a larger scale, suppose there are 100 cities on the salesman's list, the possible itineraries will be as many as 9.3×10^{157}. The problem will take many years to calculate manually, which is where optimization software steps in, automating the search for the optimal itinerary.

The example illustrated up to now is a deterministic optimization problem, that is, the airline ticket prices are known ahead of time and are assumed to be constant. Now suppose the ticket prices are not constant but are uncertain, following some distribution (e.g., a ticket from Chicago to Seattle averages \$325, but is never cheaper than \$300 and usually never exceeds \$500).[4] The same uncertainty applies to tickets for the other cities. The problem now becomes an *optimization under uncertainty*. Ad hoc and brute force approaches simply do not work. Software such as Risk Simulator can take over this optimization problem and automate the entire process seamlessly. The next section discusses the terms required in an optimization under

Seattle-Chicago-New York	$325 + $325 = **$650**
Seattle-New York-Chicago	$375 + $325 = **$700**
Chicago-Seattle-New York	$225 + $375 = **$600**
Chicago-New York-Seattle	$325 + $350 = **$675**
New York-Seattle-Chicago	$350 + $325 = **$675**
New York-Chicago-Seattle	$325 + $225 = **$550**

Additionally, say you want to visit San Antonio and Denver

For the five cities to visit (Seattle, Chicago, New York, San Antonio, and Denver) you now have:

5! = 5 x 4 x 3 x 2 x 1 = 120 possible combinations

What about 100 different cities?

100! = 100 x 99 x 98 … x 1 =
93,326,215,443,944,200,000,000,000,000,000,000,000,000,
000,000,000,000,000,000,000,000,000,000,000,000,000,
000,000,000,000,000,000,000,000,000,000,000,000,000,
000,000,000,000,000,000,000,000,000,000,000,000,000

or 9.3 X 10^{157} different combinations

FIGURE 10.3 Multiple combinations of the traveling financial planner problem.

uncertainty. Chapter 11 illustrates several additional business cases and models with step-by-step instructions.

> Optimization problems can be solved using different approaches, including the use of simplex or graphical methods, brute force, mathematically taking calculus derivatives, or using software.

THE LINGO OF OPTIMIZATION

Before embarking on solving an optimization problem, it is vital to understand the terminology of optimization—the terms used to describe certain attributes of the optimization process. These words include decision variables, constraints, and objectives.

Decision variables are quantities over which you have control; for example, the amount of a product to make, the number of dollars to allocate among different investments, or which projects to select from among a limited set. As an example, portfolio optimization analysis includes a go or no-go decision on particular projects. In addition, the dollar or percentage budget allocation across multiple projects also can be structured as decision variables.

Constraints describe relationships among decision variables that restrict the values of the decision variables. For example, a constraint might ensure that the total amount of money allocated among various investments cannot exceed a specified amount or at most one project from a certain group can be selected based on budget constraints, timing restrictions, minimum returns, or risk tolerance levels.

Objectives give a mathematical representation of the model's desired outcome, such as maximizing profit or minimizing cost, in terms of the decision variables. In financial analysis, for example, the objective may be to maximize returns while minimizing risks (maximizing the Sharpe ratio, or the returns-to-risk ratio).

Conceptually, an optimization model might look like Figure 10.4. The solution to an optimization model provides a set of values for the decision variables that optimizes (maximizes or minimizes) the associated objective. If the real business conditions were simple and the future were predictable, all data in an optimization model would be constant, making the model deterministic. In many cases, however, a deterministic optimization model cannot capture all the relevant intricacies of a practical decision-making environment. When a model's data are uncertain and can only be described

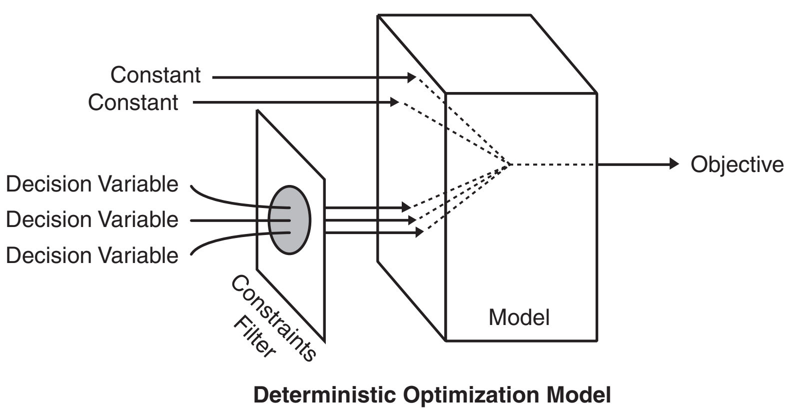


FIGURE 10.4 Visualizing a deterministic optimization.

probabilistically, the objective will have some probability distribution for any chosen set of decision variables. You can find this probability distribution by simulating the model using Risk Simulator. An optimization model under uncertainty has several additional elements, including assumptions and forecasts.

Assumptions capture the uncertainty of model data using probability distributions, whereas forecasts are the frequency distributions of possible results for the model. Forecast statistics are summary values of a forecast distribution, such as the mean, standard deviation, and variance. The optimization process controls the optimization by maximizing or minimizing the objective (see Figure 10.5).

Each optimization model has one objective, a variable that mathematically represents the model's objective in terms of the assumption and decision variables. Optimization's job is to find the optimal (minimum or maximum) value of the objective by selecting and improving different values for the decision variables. When model data are uncertain and can only be described using probability distributions, the objective itself will have some probability distribution for any set of decision variables.

> Before embarking on solving an optimization problem, the analyst first has to understand the lingo of optimization: objectives, constraints, decision variables, assumptions, and forecasts.

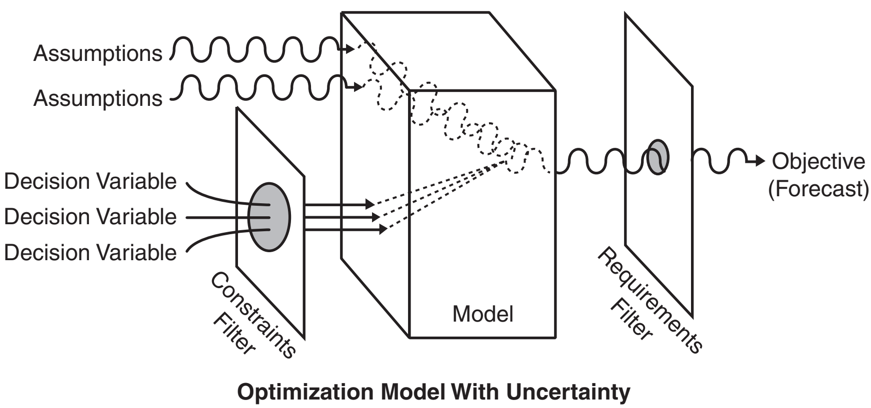


FIGURE 10.5 Visualizing a stochastic optimization.

SOLVING OPTIMIZATION GRAPHICALLY AND USING EXCEL'S SOLVER

Figure 10.6 illustrates a simple multiple constraint optimization problem solved using the graphical method. In this simple example of deterministic linear optimization with linear constraints, the graphical approach is easy to implement. However, great care should be taken when nonlinear constraints exist.[5] Sometimes, optimization models are specified incorrectly. For instance, Figure 10.7 shows problems arising with unbounded solutions (with a solution set at infinity), no feasible solution (where the constraints are too restrictive and impossible to satisfy), and multiple solutions (this is good news for management as it can choose from among several equally optimal solutions).

Figure 10.8 illustrates the same problem but solved using Excel's Solver add-in.[6] Solver is clearly a more powerful approach than the manual graphical method. This situation is especially true when multiple decision variables exist as a multidimensional graph would be required.[7] Figures 10.9 and 10.10 show the use of Solver to optimize a portfolio of projects—the former assumes an integer optimization, where projects are either a go or no-go decision, whereas the latter assumes a continuous optimization, where projects can be funded anywhere from 0 percent to 100 percent.[8]

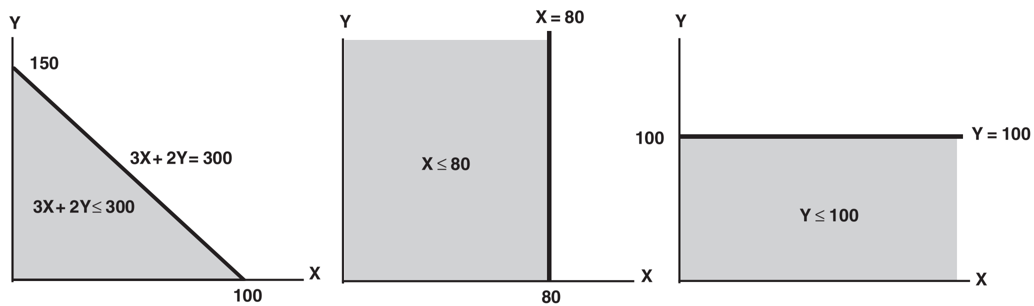
(*Text continues on page 361*)

Say there are two products X and Y being manufactured. Product X provides a \$20 profit and product Y a \$15 profit. Product X takes 3 hours to manufacture and product Y takes 2 hours to produce. In any given week, the manufacturing equipment can make both products but has a maximum capacity of 300 hours. In addition, based on market demand, management has determined that it cannot sell more than 80 units of X and 100 units of Y in a given week and prefers not to have any inventory on hand. Therefore, management has set these demand levels as the maximum output for products X and Y respectively. The issue now becomes what is the optimal production levels of both X and Y such that profits would be maximized in any given week?

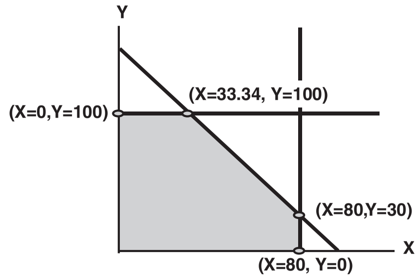
Based on the situation above, we can formulate a linear optimization routine where we have:

The Objective Function:	Max $20X + 15Y$
Subject to Constraints:	$3X + 2Y \leq 300$ $X \leq 80$ $Y \leq 100$

We can more easily visualize the constraints by plotting them out one at a time as follows:



The graph below shows the combination of all three constraints. The shaded region shows the feasible area, where all constraints are simultaneously satisfied. Hence, the optimal decision should fall within this shaded region.



We can easily calculate the intersection points of the constraints. For example, the intersection between Y = 100 and 3X + 2Y = 300 is obtained by solving the equations simultaneously. Substituting, we get 3X + 2(100) = 300. Solving yields X = 33.34 and Y = 100.

Similarly, the intersection between X = 80 and 3X + 2Y = 300 can be obtained by solving the equations simultaneously. Substituting yields 3(80) + 2Y = 300. Solving yields Y = 30 and X = 80.

The other two edges are simply intersections between the axes. Hence, when X = 80, Y = 0 for the X = 80 line and Y = 100 and X = 0 for the Y = 100 line.

From linear programming theory, one of these four intersection edges or extreme values is the optimal solution. One method is simply to substitute each of the end points into the objective function and see which solution set provides the highest profit level.

Using the objective function where Profit = 20X + 15Y and substituting each of the extreme value sets:

When X = 0 and Y = 100:	Profit = $20 (0) + $15 (100) = $1,500
When X = 33.34 and Y = 100:	Profit = $20 (33.34) + $15 (100) = $2,167
When X = 80 and Y = 30:	Profit = $20 (80) + $15 (30) = $2,050
When X = 80 and Y = 0:	Profit = $20 (80) + $15 (0) = $1,600

Here, we see that when X = 33.34 and Y = 100, the profit function is maximized. We can also further verify this by using any combinations of X and Y within the feasible (shaded) area above. For instance, X =10 and Y =10 is a combination that is feasible but their profit outcome is only $20 (10) + $15 (10) = $350. We can calculate infinite combinations of X and Y sets but the optimal combination is always going to be at extreme value edges.

We can easily verify which extreme value will be the optimal solution set by drawing the objective function line. If we set the objective function to be:

20X + 15Y = 0	we get X = 20, Y = 15
20X + 15Y = 1200	we get X = 60, Y = 80

If we keep shifting the profit function upwards to the right, we will keep intersecting with the extreme value edges. The edge that provides the highest profit function is the optimal solution set.

In our example, point B is the optimal solution, which was verified by our calculations above, where X = 33.34 and Y = 100.

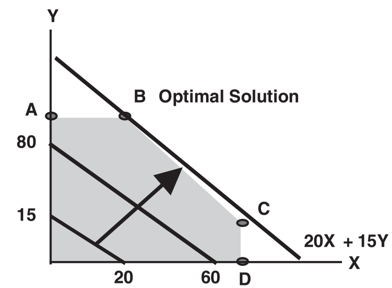
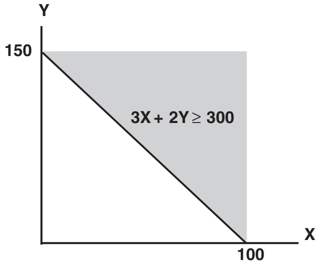


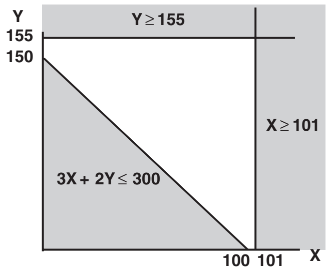
FIGURE 10.6 Solving optimization using linear programming.

There could be potential problems when dealing with linear programming. The three most frequently occurring problems include: Unbounded Solutions, No Feasible Solutions, and Multiple Optimal Solutions.



Unbounded Solutions

For instance, if the only constraint was such that $3X + 2Y \geq 300$, we have an unbounded problem. This means the machine can keep working greater than 300 hours without stop. Hence, optimally, in order to generate the most amount of profit, we would keep making products X and Y up to an infinite level. This is essentially management heaven, to produce as much as possible without any budgetary or resource constraints. Obviously, if this is the case, we should assume that the problem has not been defined correctly and perhaps an error has occurred in our mathematical models.



No Feasible Solution

Now suppose we have the following constraints:

$$3X + 2Y \leq 300$$
$$X \geq 101$$
$$Y \geq 155$$

There exists no area where all constraints are binding simultaneously. In essence, any solution generated will by definition not be feasible since there will always be a constraint that is violated. Given a situation like this, it may be that the problem has been framed incorrectly or that we may have to request that management loosen some of its tight constraints since based on its expectations, the project is just not doable. Additional resources are required (greater than 300 hours by purchasing additional machines or hiring more workers) or that the minimum required production levels (155 and 101) be reduced.

Multiple Solutions

Here, we have two extreme values (B and C) that intersect the profit objective function. Both these solution sets are optimal. This is good news for management since it has the option of choosing either combination of X and Y production levels. Other qualitative factors may be utilized on top of quantitative analytical results.

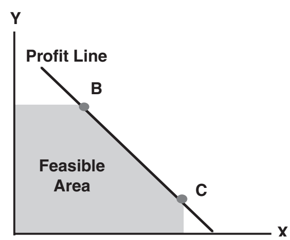


FIGURE 10.7 Potential problems of linear programming.

Using the same previous problem, where we have the following:

The Objective Function: Max $20X + 15Y$

Subject to Constraints: $3X + 2Y \leq 300$

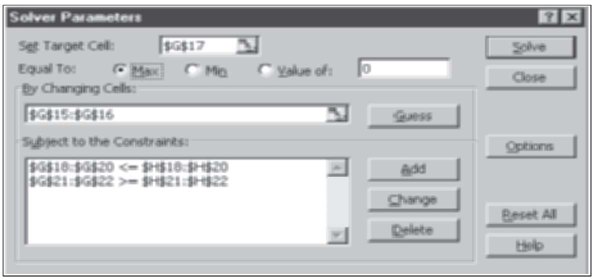
$X \leq 80$

$Y \leq 100$

We can utilize Excel's Solver add-in to provide a quick analytical solution.

	F	G	H
15	X		
16	Y		
17	Profit	=20*G15+15*G16	
18	Constraints	=3*G15+2*G16	300
19		=G15	80
20		=G16	100
21	Nonnegative	=G15	0
22		=G16	0

First, we need to set up the spreadsheet model. We have an X and Y variable which is to be solved. Next, we have the profit objective function in cell G17 and the constraints in cells G18 through H22. In order for Solver to perform the calculation, we needed to include two additional requirements, the nonnegative constraints, where we are setting both X and Y to be positive values only. Negative values of production are impossible. Cells H18 to H22 are the target values for the constraints. We then start Solver by clicking on Tools and Solver. (If Solver is not available, you may have to first add it in by clicking on Tools/Add-Ins and selecting Solver. Then, go back to Tools/Solver to run the program).



Set the profit calculation as the target cell ($G$17) and select maximization. Set the X and Y unknowns as the cells to change ($G$15:$G$16). Next, click on Add to add the constraints. The constraints could be added one at a time or in a batch group. Add $G$18:$G$20 to be less than or equal to $H$18:$H$20. Then, add in the nonnegative constraints where $G$21:$G$22 is greater than or equal to zero ($H$21:$H$22).

If we let Solver calculate the results, we would obtain the following, where the optimal solution set is when:

X	33.33	
Y	100	
Profit	$2,167	
Constraints	300	300
	33.33	80
	100	100
Nonnegative	33.33	0
	100	0

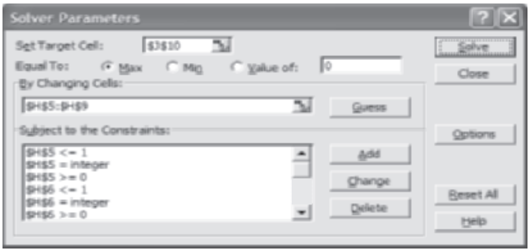
FIGURE 10.8 Using Excel's Solver in linear programming.

Integer Portfolio Optimization and Integer Linear Programming

	Cost	Return	Risk	Return-Risk Ratio	Allocation	Weighted Cost	Risk Return	Weighted Risk
Project A	$500,000	19%	32%	0.594	0%	$0	0.000	0%
Project B	$625,000	23%	39%	0.590	0%	$0	0.000	0%
Project C	$345,000	15%	22%	0.682	100%	$345,000	0.682	22%
Project D	$290,000	16%	29%	0.552	0%	$0	0.000	0%
Project E	$450,000	17%	25%	0.680	100%	$450,000	0.680	25%
					Sum	$795,000	1.362	47%

Budget Constraint $1,000,000

Each project must be between 10% and 50% allocated in funds



Suppose you have 5 projects you wish to allocate a fixed budget of $500,000 (this is your constraint) among, such that you will maximize the return to risk ratio (this is the objective function) subject to the requirements that each of these projects can be allocated anywhere between 10% and 50% of its total cost. You cannot allocate more than 50% of the cost of a project since you are only in the beginning stages of development while at least 10% of the project should be funded since all five projects have been previously found to be financially feasible. Using Excel's Solver add-in (use Tools/Add-Ins/Solver and then Tools/Solver) we calculate the optimal weights that will maximize the return to risk ratio.

Target cell is the objective function, which in this case is the total return to risk ratio weighted by each project, which is to be maximized. Next, add additional constraints including the budget constraint where the total cost allocated in the portfolio is ≤ the budget constraint. In addition, for each project weight, set them to be ≥ 0 and ≤ 1 as well as weight as integers. This is essentially the difference between the prior linear programming and optimization routine which allows fractional projects to be executed while in integer linear programming, projects are either chosen (1.0) or not (0.0) and nothing in between is allowed (integer constraint).

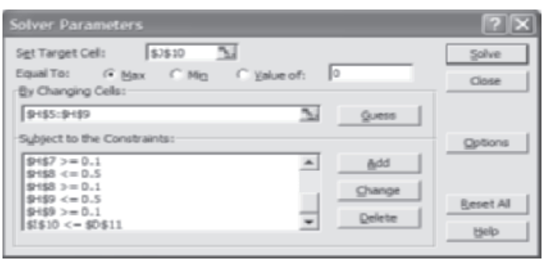
FIGURE 10.9 Excel Solver on integer linear programming.

Portfolio Optimization and Linear Programming

	Cost	Return	Risk	Return-Risk Ratio	Allocation	Weighted Cost	Total Risk-Return	Weighted Risk
Project A	$500,000	19%	32%	0.594	10%	$50,000	0.059	3%
Project B	$625,000	23%	39%	0.590	10%	$62,500	0.059	4%
Project C	$345,000	15%	22%	0.682	50%	$172,500	0.341	11%
Project D	$290,000	16%	29%	0.552	50%	$145,000	0.276	15%
Project E	$450,000	17%	25%	0.680	16%	$70,000	0.106	4%
					Sum	$500,000	0.841	36%

Budget Constraint $500,000

Each project must be between 10% and 50% allocated in funds



Suppose you have 5 projects you wish to allocate a fixed budget of $500,000 (this is your constraint) among, such that you will maximize the return to risk ratio (this is the objective function) subject to the requirements that each of these projects can be allocated anywhere between 10% and 50% of its total cost. You cannot allocate more than 50% of the cost of a project since you are only in the beginning stages of development while at least 10% of the project should be funded since all five projects have been previously found to be financially feasible. Using Excel's Solver add-in (use Tools/Add-Ins/Solver and then Tools/Solver) we calculate the optimal weights that will maximize the return to risk ratio.

Target cell is the objective function, which in this case is the total return to risk ratio weighted by each project, which is to be maximized. Next, add additional constraints including the budget constraint where the total cost allocated in the portfolio is ≤ the budget constraint. In addition, for each project weight, set them to be ≥ 0.1 and ≤ 0.5.

FIGURE 10.10 Excel Solver on continuous linear programming.

QUESTIONS

1. What is the difference between deterministic optimization and optimization under uncertainty?
2. Define then compare and contrast each of the following:
   a. Objective
   b. Constraint
   c. Decision variable
3. Explain what some of the problems are in a graphical linear programming approach and if they can be easily solved.
4. What are some of the approaches to solve an optimization problem? List each approach as well as its corresponding pros and cons.

CHAPTER 11

Optimization Under Uncertainty

This chapter looks at the optimization process and methodologies in more detail as it pertains to using Risk Simulator. These methodologies include the use of continuous versus discrete integer optimization, as well as static versus dynamic and stochastic optimizations. The chapter then proceeds with two example optimization models to illustrate how the optimization process works. The first is the application of *continuous* optimization under uncertainty for a simple project selection model, where the idea is to allocate 100 percent of an individual's investment among several different asset classes (e.g., different types of mutual funds or investment styles: growth, value, aggressive growth, income, global, index, contrarian, momentum, and so forth). The second project deals with *discrete integer* optimization, where the idea is to look at several competing and nonmutually exclusive project choices, each with a different return, risk, and cost profile. The job of the analyst here is to find the best combination of projects that will satisfy the firm's budget constraints while maximizing the portfolio's total value.

OPTIMIZATION PROCEDURES

Many algorithms exist to run optimization and many different procedures exist when optimization is coupled with Monte Carlo simulation. In Risk Simulator, there are three distinct optimization procedures and optimization types as well as different decision variable types. For instance, Risk Simulator can handle *continuous decision variables* (1.2535, 0.2215, and so forth), *integers decision variables* (1, 2, 3, 4), *binary decision variables* (1 and 0 for go and no-go decisions), and *mixed decision variables* (both integers and continuous variables). On top of that, Risk Simulator can handle *linear optimization* (i.e., when both the objective and constraints are all linear equations and functions) and *nonlinear optimizations* (i.e., when the objective and constraints are a mixture of linear and nonlinear functions and equations).

As far as the optimization process is concerned, Risk Simulator can be used to run a *discrete optimization*, that is, an optimization that is run on a discrete or static model, where no simulations are run. In other words, all the inputs in the model are static and unchanging. This optimization type is applicable when the model is assumed to be known and no uncertainties exist. Also, a discrete optimization can first be run to determine the optimal portfolio and its corresponding optimal allocation of decision variables before more advanced optimization procedures are applied. For instance, before running a stochastic optimization problem, a discrete optimization is first run to determine if solutions to the optimization problem exist before a more protracted analysis is performed.

Next, *dynamic optimization* is applied when Monte Carlo simulation is used together with optimization. Another name for such a procedure is *simulation-optimization*; that is, a simulation is first run, then the results of the simulation are applied in the Excel model, and then an optimization is applied to the simulated values. In other words, a simulation is run for N trials, and then an optimization process is run for M iterations until the optimal results are obtained or an infeasible set is found. Using Risk Simulator's optimization module, you can choose which forecast and assumption statistics to use and replace in the model after the simulation is run. Then, these forecast statistics can be applied in the optimization process. This approach is useful when you have a large model with many interacting assumptions and forecasts, and when some of the forecast statistics are required in the optimization. For example, if the standard deviation of an assumption or forecast is required in the optimization model (e.g., computing the Sharpe ratio in asset allocation and optimization problems where we have mean divided by standard deviation of the portfolio), then this approach should be used.

The *stochastic optimization* process, in contrast, is similar to the dynamic optimization procedure with the exception that the entire dynamic optimization process is repeated T times; that is, a simulation with N trials is run, and then an optimization is run with M iterations to obtain the optimal results. Then the process is replicated T times. The results will be a forecast chart of each decision variable with T values. In other words, a simulation is run and the forecast or assumption statistics are used in the optimization model to find the optimal allocation of decision variables. Then, another simulation is run, generating different forecast statistics, and these new updated values are then optimized, and so forth. Hence, the final decision variables will each have their own forecast chart, indicating the range of the optimal decision variables. For instance, instead of obtaining single-point estimates in the dynamic optimization procedure, you can now obtain a distribution of the decision variables, hence, a range of optimal values for each decision variable, also known as a stochastic optimization.

Finally, an efficient frontier optimization procedure applies the concepts of marginal increments and shadow pricing in optimization; that is, what would happen to the results of the optimization if one of the constraints were relaxed slightly? Say for instance, if the budget constraint is set at $1 million. What would happen to the portfolio's outcome and optimal decisions if the constraint were now $1.5 million, or $2 million, and so forth? This is the concept of the Markowitz efficient frontier in investment finance, where if the portfolio standard deviation is allowed to increase slightly, what additional returns will the portfolio generate? This process is similar to the dynamic optimization process with the exception that *one* of the constraints is allowed to change, and with each change, the simulation and optimization process is run. This process is best applied manually using Risk Simulator. Run a dynamic or stochastic optimization, then rerun another optimization with a new constraint, and repeat that procedure several times. This manual process is important, as by changing the constraint, the analyst can determine if the results are similar or different, and hence, whether it is worthy of any additional analysis, or to determine how far a marginal increase in the constraint should be to obtain a significant change in the objective and decision variables. This is done by comparing the forecast distribution of each decision variable after running a stochastic optimization.

One item is worthy of consideration. Other software products exist that supposedly perform stochastic optimization, but, in fact, they do not. For instance, after a simulation is run, then *one* iteration of the optimization process is generated, and then another simulation is run, then the *second* optimization iteration is generated, and so forth. This process is simply a waste of time and resources; that is, in optimization, the model is put through a rigorous set of algorithms, where multiple iterations (ranging from several to thousands of iterations) are required to obtain the optimal results. Hence, generating *one* iteration at a time is a waste of time and resources. The same portfolio can be solved using Risk Simulator in under a minute as compared to multiple hours using such a backward approach. Also, such a simulation-optimization approach will typically yield bad results and is not a stochastic optimization approach. Be extremely careful of such methodologies when applying optimization to your models.

The following are two example optimization problems. One uses continuous decision variables while the other uses discrete integer decision variables. In either model, you can apply discrete optimization, dynamic optimization, stochastic optimization, or even manually generate efficient frontiers with shadow pricing. Any of these approaches can be used for these two examples. Therefore, for simplicity, only the model setup is illustrated and it is up to the user to decide which optimization process to run. Also, the continuous decision variable example uses the nonlinear optimization approach (because the portfolio risk computed is a nonlinear function, and the objective is a nonlinear function of portfolio returns divided by portfolio risks) while the second example of an integer optimization is an example of a linear optimization model (its objective and all of its constraints are linear). Therefore, these two examples encapsulate all of the procedures aforementioned.

Note that the examples in this chapter use Risk Simulator's optimization module, where some of the algorithms apply Excel's Solver add-in. You will need Solver installed to run these models. Install Solver by clicking on *Tools* and *Add-Ins* in Excel, and selecting *Solver*. Depending how Excel was initially installed on your computer, you may or may not be required to insert the Microsoft Office installation CD to continue.

CONTINUOUS OPTIMIZATION

Figure 11.1 illustrates the sample continuous optimization model. The example here uses the *Continuous Optimization* file found on *Start* | *Programs* | *Real Options Valuation* | *Risk Simulator* | *Examples*. In this example, there are 10 distinct asset classes (e.g., different types of mutual funds, stocks, or assets) where the idea is to most efficiently and effectively allocate the portfolio holdings such that the best *bang for the buck* is obtained; that is, to generate the best portfolio returns possible given the risks inherent in each asset class. In order to truly understand the concept of optimization, we must delve more deeply into this sample model to see how the optimization process can best be applied.

The model shows the 10 asset classes and each asset class has its own set of annualized returns and annualized volatilities. These return and risk measures are annualized values such that they can be consistently compared across different asset classes. Returns are computed using the geometric average of the relative returns while the risks are computed using the logarithmic relative stock returns approach. See the appendix to this chapter for details on computing the annualized volatility and annualized returns on a stock or asset class.

The Allocation Weights in column E hold the decision variables, which are the variables that need to be tweaked and tested such that the total weight is constrained at 100 percent (cell E17). Typically, to start the optimization, we will set these cells to a uniform value, where in this case, cells E6 to E15 are set at 10 percent each. In addition, each decision variable may have specific restrictions in its allowed range. In this example, the lower and upper allocations allowed are 5 percent and 35 percent, as seen in columns F and G. This means that each asset class may have its own allocation boundaries. Next, column H shows the return to risk ratio, which is simply the return percentage divided by the risk percentage, where the higher this

ASSET ALLOCATION OPTIMIZATION MODEL

Asset Class Description	Annualized Returns	Volatility Risk	Allocation Weights	Required Minimum Allocation	Required Maximum Allocation	Return to Risk Ratio	Returns Ranking (Hi-Lo)	Risk Ranking (Lo-Hi)	Return to Risk Ranking (Hi-Lo)	Allocation Ranking (Hi-Lo)
Asset Class 1	10.54%	12.36%	10.00%	5.00%	35.00%	0.8524	9	2	7	1
Asset Class 2	11.25%	16.23%	10.00%	5.00%	35.00%	0.6929	7	8	10	1
Asset Class 3	11.84%	15.64%	10.00%	5.00%	35.00%	0.7570	6	7	9	1
Asset Class 4	10.64%	12.35%	10.00%	5.00%	35.00%	0.8615	8	1	5	1
Asset Class 5	13.25%	13.28%	10.00%	5.00%	35.00%	0.9977	5	4	2	1
Asset Class 6	14.21%	14.39%	10.00%	5.00%	35.00%	0.9875	3	6	3	1
Asset Class 7	15.53%	14.25%	10.00%	5.00%	35.00%	1.0898	1	5	1	1
Asset Class 8	14.95%	16.44%	10.00%	5.00%	35.00%	0.9094	2	9	4	1
Asset Class 9	14.16%	16.50%	10.00%	5.00%	35.00%	0.8584	4	10	6	1
Asset Class 10	10.06%	12.50%	10.00%	5.00%	35.00%	0.8045	10	3	8	1
Portfolio Total	*12.6419%*	*4.58%*	*100.00%*							
Return to Risk Ratio	*2.7596*									

Specifications of the optimization model:

Objective:	*Maximize Return to Risk Ratio (C18)*
Decision Variables:	*Allocation Weights (E6:E15)*
Restrictions on Decision Variables:	*Minimum and Maximum Required (F6:G15)*
Constraints:	*Portfolio Total Allocation Weights 100% (E17 is set to 100%)*

Additional specifications:

1. One can always maximize portfolio total returns or minimize the portfolio total risk.
2. Incorporate Monte Carlo simulation in the model by simulating the returns and volatility of each asset class and apply Simulation-Optimization techniques.
3. The portfolio can be optimized as is without simulation using Static Optimization techniques.

FIGURE 11.1 Continuous optimization model.

value, the higher the *bang for the buck*. The remaining model shows the individual asset class rankings by returns, risk, return to risk ratio, and allocation. In other words, these rankings show at a glance which asset class has the lowest risk, or the highest return, and so forth.

The portfolio's total returns in cell C17 are *SUMPRODUCT(C6:C15, E6:E15)*, that is, the sum of the allocation weights multiplied by the annualized returns for each asset class. In other words, we have $R_P = \omega_1 R_1 + \omega_2 R_2 + \omega_3 R_3 + \ldots \omega_{10} R_{10}$, where R_P is the return on the portfolio, R_i are the individual returns on the projects, and ω_i are the respective weights or capital allocation across each project.

In addition, the portfolio's diversified risk in cell D17 is computed by taking

$$\sigma_P = \sqrt{\sum_{i=1}^{n} \omega_i^2 \sigma_i^2 + \sum_{i=1}^{n} \sum_{j=1}^{m} 2\omega_i \omega_j \rho_{i,j} \sigma_i \sigma_j}$$

Here, ρ_{ij} are the respective cross correlations between the asset classes. Hence, if the cross correlations are negative, there are risk diversification effects, and the portfolio risk decreases. However, to simplify the computations here, we assume zero correlations among the asset classes through this portfolio risk computation, but assume the correlations when applying simulation on the returns as will be seen later. Therefore, instead of applying static correlations among these different asset returns, we apply the correlations in the simulation assumptions themselves, creating a more dynamic relationship among the simulated return values.

Finally, the return to risk ratio or Sharpe ratio is computed for the portfolio. This value is seen in cell C18 and represents the objective to be maximized in this optimization exercise. To summarize, we have the following specifications in this example model:

Objective:	*Maximize Return to Risk Ratio (C18)*
Decision Variables:	*Allocation Weights (E6:E15)*
Restrictions on Decision Variables:	*Minimum and Maximum Required (F6:G15)*
Constraints:	*Total Allocation Weights Sum to 100% (E17)*

Procedure

Use the following procedure to run an optimization analysis:

1. Open the example file (*Continuous Optimization*) and start a new profile by clicking on ***Simulation* | *New Profile*** and provide it a name.

2. The first step in optimization is to set the decision variables. Select cell E6 and set the first decision variable (***Simulation* | *Optimization* | *Set Decision***) and click on the link icon to select the name cell (B6), as well as the lower bound and upper bound values at cells F6 and G6. Then, using Risk Simulator copy, copy this cell E6 decision variable and paste the decision variable to the remaining cells in E7 to E15.
3. The second step in optimization is to set the constraint. There is only one constraint here, that is, the total allocation in the portfolio must sum to 100%. So, click on ***Simulation* | *Optimization* | *Constraints . . .*** and select ***ADD*** to add a new constraint. Then, select the cell E17 and make it equal (=) to 100%. Click OK when done.
4. The final step in optimization is to set the objective function and start the optimization by selecting the objective cell C18 and ***Simulation* | *Optimization* | *Set Objective*** and then run the optimization by selecting ***Simulation* | *Optimization* | *Run Optimization*** choosing the optimization of choice (Static Optimization, Dynamic Optimization, or Stochastic Optimization). To get started, select ***Static Optimization***. Check to make sure the objective cell is set for C18 and select ***Maximize***. You can now review the decision variables and constraints if required, or click OK to run the static optimization.
5. Once the optimization is complete, you may select ***Revert*** to revert back to the original values of the decision variables as well as the objective, or select ***Replace*** to apply the optimized decision variables. Typically, Replace is chosen after the optimization is done.

Figure 11.2 shows the screen shots of the preceding procedural steps. You can add simulation assumptions on the model's returns and risk (columns C and D) and apply the dynamic optimization and stochastic optimization for additional practice.

Results Interpretation

The optimization's final results are shown in Figure 11.3, where the optimal allocation of assets for the portfolio is seen in cells E6:E15. Given the restrictions of each asset fluctuating between 5 percent and 35 percent, and where the sum of the allocation must equal 100 percent, the allocation that maximizes the return to risk ratio is seen in Figure 11.3.

A few important things must be noted when reviewing the results and optimization procedures performed thus far:

- The correct way to run the optimization is to maximize the bang for the buck or returns to risk Sharpe ratio as we have done.

FIGURE 11.2 Running continuous optimization in Risk Simulator.

ASSET ALLOCATION OPTIMIZATION MODEL

Asset Class Description	Annualized Returns	Volatility Risk	Allocation Weights	Required Minimum Allocation	Required Maximum Allocation	Return to Risk Ratio	Returns Ranking (Hi-Lo)	Risk Ranking (Lo-Hi)	Return to Risk Ranking (Hi-Lo)	Allocation Ranking (Hi-Lo)
Asset Class 1	10.54%	12.36%	11.09%	5.00%	35.00%	0.8524	9	2	7	4
Asset Class 2	11.25%	16.23%	6.87%	5.00%	35.00%	0.6929	7	8	10	10
Asset Class 3	11.84%	15.64%	7.78%	5.00%	35.00%	0.7570	6	7	9	9
Asset Class 4	10.64%	12.35%	11.22%	5.00%	35.00%	0.8615	8	1	5	3
Asset Class 5	13.25%	13.28%	12.08%	5.00%	35.00%	0.9977	5	4	2	2
Asset Class 6	14.21%	14.39%	11.04%	5.00%	35.00%	0.9875	3	6	3	5
Asset Class 7	15.53%	14.25%	12.30%	5.00%	35.00%	1.0898	1	5	1	1
Asset Class 8	14.95%	16.44%	8.90%	5.00%	35.00%	0.9094	2	9	4	7
Asset Class 9	14.16%	16.50%	8.37%	5.00%	35.00%	0.8584	4	10	6	8
Asset Class 10	10.06%	12.50%	10.35%	5.00%	35.00%	0.8045	10	3	8	6
Portfolio Total	***12.6920%***	***4.52%***	***100.00%***							
Return to Risk Ratio	***2.8091***									

Specifications of the optimization model:

Objective: *Maximize Return to Risk Ratio (C18)*
Decision Variables: *Allocation Weights (E6:E15)*
Restrictions on Decision Variables: *Minimum and Maximum Required (F6:G15)*
Constraints: *Portfolio Total Allocation Weights 100% (E17 is set to 100%)*

Additional specifications:

1. One can always maximize portfolio total returns or minimize the portfolio total risk.
2. Incorporate Monte Carlo simulation in the model by simulating the returns and volatility of each asset class and apply Simulation-Optimization techniques.
3. The portfolio can be optimized as is without simulation using Static Optimization techniques.

FIGURE 11.3 Continuous optimization results.

- If instead we maximized the total portfolio returns, the optimal allocation result is trivial and does not require optimization to obtain; that is, simply allocate 5 percent (the minimum allowed) to the lowest eight assets, 35 percent (the maximum allowed) to the highest returning asset, and the remaining (25 percent) to the second-best returns asset. Optimization is not required. However, when allocating the portfolio this way, the risk is a lot higher as compared to when maximizing the returns to risk ratio, although the portfolio returns by themselves are higher.
- In contrast, one can minimize the total portfolio risk, but the returns will now be less.

Table 11.1 illustrates the results from the three different objectives being optimized. From the table, the best approach is to maximize the returns to risk ratio, that is, for the same amount of risk, this allocation provides the highest amount of return. Conversely, for the same amount of return, this allocation provides the lowest amount of risk possible. This approach of *bang for the buck* or returns to risk ratio is the cornerstone of the Markowitz efficient frontier in modern portfolio theory. That is, if we constrain the total portfolio risk levels and successively increase them over time, we will obtain several efficient portfolio allocations for different risk characteristics. Thus, different efficient portfolio allocations can be obtained for different individuals with different risk preferences.

DISCRETE INTEGER OPTIMIZATION

Sometimes, the decision variables are not continuous but discrete integers (e.g., 1, 2, 3) or binary (e.g., 0 and 1). We can use such binary decision variables as on-off switches or go/no-go decisions. Figure 11.4 illustrates a project selection model where there are 20 projects listed. The example here uses

TABLE 11.1 Optimization Results

Objective	Portfolio Returns (%)	Portfolio Risk (%)	Portfolio Returns to Risk Ratio
Maximize Returns to Risk Ratio	12.69	4.52	2.8091
Maximize Returns	13.97	6.77	2.0636
Minimize Risk	12.38	4.46	2.7754

Project Name	ENPV	NPV	Cost	Risk	Return to Risk Ratio	Profitability Index	Selection
Project 1	$458.00	$150.76	$1,732.44	12.00%	3816.67	1.09	1
Project 2	$1,954.00	$245.00	$859.00	98.00%	1993.88	1.29	1
Project 3	$1,599.00	$458.00	$1,845.00	97.00%	1648.45	1.25	1
Project 4	$2,251.00	$529.00	$1,645.00	45.00%	5002.22	1.32	1
Project 5	$849.00	$564.00	$458.00	109.00%	778.90	2.23	1
Project 6	$758.00	$135.00	$52.00	74.00%	1024.32	3.60	1
Project 7	$2,845.00	$311.00	$758.00	198.00%	1436.87	1.41	1
Project 8	$1,235.00	$754.00	$115.00	75.00%	1646.67	7.56	1
Project 9	$1,945.00	$198.00	$125.00	108.00%	1800.93	2.58	1
Project 10	$2,250.00	$785.00	$458.00	85.00%	2647.06	2.71	1
Project 11	$549.00	$35.00	$45.00	48.00%	1143.75	1.78	1
Project 12	$525.00	$75.00	$105.00	59.00%	889.83	1.71	1
Project 13	$516.00	$451.00	$48.00	28.00%	1842.86	10.40	1
Project 14	$499.00	$458.00	$351.00	94.00%	530.85	2.30	1
Project 15	$859.00	$125.00	$421.00	65.00%	1321.54	1.30	1
Project 16	$884.00	$458.00	$124.00	39.00%	2266.67	4.69	1
Project 17	$956.00	$124.00	$521.00	154.00%	620.78	1.24	1
Project 18	$854.00	$164.00	$512.00	210.00%	406.67	1.32	1
Project 19	$195.00	$45.00	$5.00	12.00%	1625.00	10.00	1
Project 20	$210.00	$85.00	$21.00	10.00%	2100.00	5.05	1
Total	$22,191.00		$10,200.44	438%			20
Goal:	MAX		<=$5000				<=10
Sharpe Ratio	5071.41						

FIGURE 11.4 Discrete integer optimization model.

the ***Discrete Optimization*** file found on ***Start* | *Programs* | *Real Options Valuation* | *Risk Simulator* | *Examples***. Each project, like before, has its own returns (ENPV and NPV for expanded net present value and net present value—the ENPV is simply the NPV plus any strategic real options values), costs of implementation, risks, and so forth. If required, this model can be modified to include required full-time equivalences (FTE) and other resources of various functions, and additional constraints can be set on these additional resources. The inputs into this model are typically linked from other spreadsheet models. For instance, each project will have its own discounted cash flow or returns on investment model. The application here is to maximize the portfolio's Sharpe ratio subject to some budget allocation. Many other versions of this model can be created, for instance, maximizing the portfolio returns, or minimizing the risks, or add additional constraints where the total number of projects chosen cannot exceed 10, and so forth and so on. All of these items can be run using this existing model.

Procedure

Use the following procedure to set up and run the optimization:

1. Open the example file (*Discrete Optimization*) and start a new profile by clicking on ***Simulation* | *New Profile*** and provide it a name.
2. The first step in optimization is to set up the decision variables. Set the first decision variable by selecting cell J4, and select ***Simulation* | *Optimization* | *Set Decision***, click on the link icon to select the name cell (B4), and select the ***Binary*** variable. Then, using Risk Simulator copy, copy this J4 decision variable cell and paste the decision variable to the remaining cells in J5 to J23.
3. The second step in optimization is to set the constraint. There are two constraints here, that is, the total budget allocation in the portfolio must be less than $5,000 and the total number of projects must not exceed 10. So, click on ***Simulation* | *Optimization* | *Constraints . . .*** and select ***ADD*** to add a new constraint. Then, select the cell E25 and make it less than or equal (≤) to 5,000. Repeat by setting cell J25 ≤ 10.
4. The final step in optimization is to set the objective function and start the optimization by selecting cell C27 and selecting ***Simulation* | *Optimization* | *Set Objective*** and then run the optimization (***Simulation* | *Optimization* | *Run Optimization***) and choosing the optimization of choice (Static Optimization, Dynamic Optimization, or Stochastic Optimization). To get started, select ***Static Optimization***. Check to make sure that the objective cell is C27 and select ***Maximize***. You can now review the decision variables and constraints if required, or click OK to run the static optimization.

Figure 11.5 shows the screen shots of the foregoing procedural steps. You can add simulation assumptions on the model's ENPV and Risk (columns C and F) and apply the dynamic optimization and stochastic optimization for additional practice.

Results Interpretation

In contrast, one can always maximize total revenues, but as before, this process is trivial and simply involves choosing the highest returning project and going down the list until you run out of money or exceed the budget constraint. Doing so will yield theoretically undesirable projects, as the highest yielding projects typically hold higher risks. Now, if desired, you can replicate the optimization using a stochastic or dynamic optimization by adding in assumptions in the ENPV and Risk values.

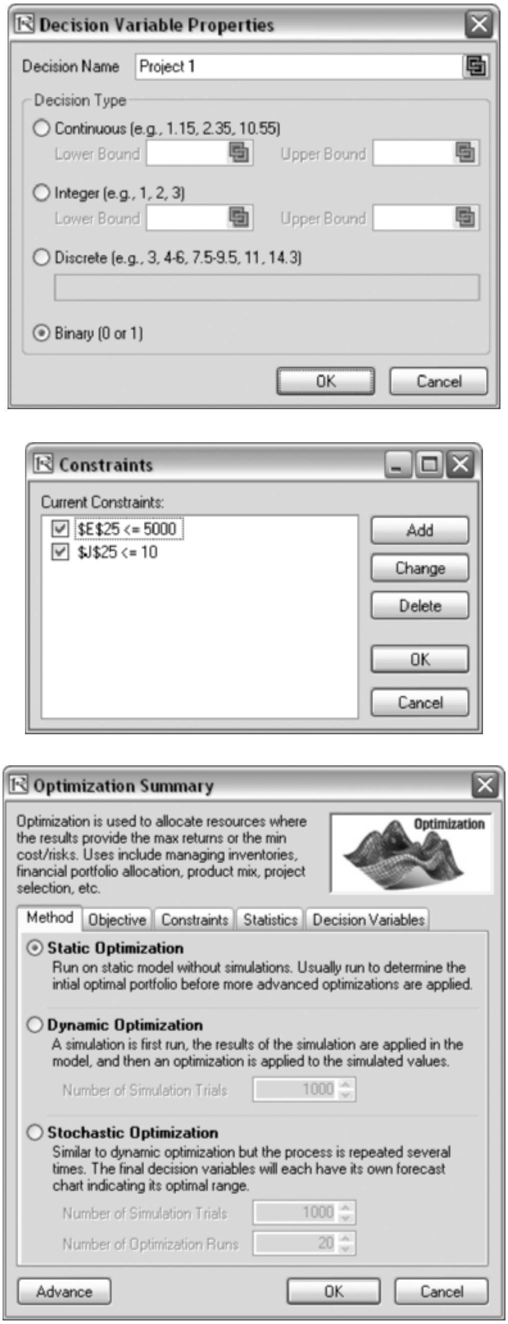


FIGURE 11.5 Running discrete integer optimization in Risk Simulator.

APPENDIX—COMPUTING ANNUALIZED RETURNS AND RISK FOR PORTFOLIO OPTIMIZATION

Figure 11.6 illustrates a quick example using Microsoft's historical stock prices for computing the annualized return and annualized volatility risk. It shows the stock prices for Microsoft downloaded from Yahoo! Finance, a publicly available free resource (visit http://finance.yahoo.com and enter a stock symbol, e.g., MSFT for Microsoft, click on *Quotes: Historical Prices*, select *Weekly*, and select the period of interest to download the data to a spreadsheet for analysis). The data in columns A and B are downloaded from Yahoo.

The formula in cell D3 is simply *LN(B3/B4)* to compute the natural logarithmic value of the relative returns week after week, and is copied down the entire column. The formula in cell E3 is *STDEV(D3:D54)*SQRT(52)* which computes the annualized (by multiplying the square root of the number of weeks in a year) volatility (by taking the standard deviation of the entire 52 weeks of the year 2004 data). The formula in cell E3 is then copied down the entire column to compute a moving window of annualized volatilities. The volatility used is this example is the average of a 52-week moving window, which covers 2 years of data; that is, cell M8's formula is *AVERAGE(E3:E54)*, where cell E54 has the following formula: *STDEV(D54:D105)*SQRT(52)*, and, of course, row 105 is January 2003. This means that the 52-week moving window captures the average volatility over a 2-year period and smoothes the volatility such that infrequent but extreme spikes will not dominate the volatility computation. Of course, a median volatility should also be computed. If the median is far off from the average, the distribution of volatilities is skewed and the median should be used; otherwise, the average should be used. Finally, these 52 volatilities can be fed into Monte Carlo simulation, using the Risk Simulator software's custom distribution to run a nonparametric simulation or to perform a data fitting procedure to find the best-fitting distribution to simulate.

In contrast, we can compute the annualized returns either using the arithmetic average method or the geometric average method. Cell G3 computes the absolute percentage return for the week where the formula for the cell is *(B3–B4)/B4*, and the formula is copied down the entire column. Then, the moving average window is computed in cell H3 as *AVERAGE(G3:G54)*52*, where the average weekly returns are obtained and annualized by multiplying them with 52, the number of weeks in a year. Note that averages are additive and can be multiplied directly by the number of weeks in a year versus volatility, which is not additive. Only volatility squared is additive, which means that the periodic volatility computed previously needs to be multiplied by the square root of 52. The arithmetic average return in cell M14 is, hence, the average of a 52-week period of the

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	**Historical Data**			**Volatility Computations**		**Returns Computations**							
2	Week	Closing Price		LN Relative Returns	Moving Average Volatilities	Relative Returns	Absolute Returns	Moving Average Absolute Returns	Moving Average Geometric Returns				
3	27-Dec-04	26.64		-0.0108	17.87%	0.9892	-1.08%	10.04%	7.69%				
4	20-Dec-04	26.93		0.0019	17.84%	1.0019	0.19%	11.98%	9.55%				
5	13-Dec-04	26.88		-0.0045	17.85%	0.9956	-0.44%	11.27%	10.22%				
6	6-Dec-04	27.00		-0.0055	18.00%	0.9945	-0.55%	14.36%	10.14%	***One-Year Annualized Volatility Analysis***			
7	29-Nov-04	27.15		0.0235	18.13%	1.0238	2.38%	17.50%	13.31%				
8	22-Nov-04	26.52		-0.0098	18.03%	0.9903	-0.97%	16.17%	13.52%	Average Annualized Volatility			**21.89%**
9	15-Nov-04	26.78		-0.0011	18.10%	0.9989	-0.11%	19.56%	15.54%	Median Annualized Volatility			**22.30%**
10	8-Nov-04	26.81		0.0223	18.20%	1.0225	2.25%	18.13%	18.05%				
11	1-Nov-04	26.22		0.0468	18.28%	1.0480	4.80%	13.56%	14.26%				
12	25-Oct-04	25.02		0.0084	17.71%	1.0085	0.85%	8.63%	7.21%	***One-Year Annualized Returns Analysis***			
13	18-Oct-04	24.81		-0.0092	17.80%	0.9908	-0.92%	6.02%	6.24%				
14	11-Oct-04	25.04		0.0000	19.68%	1.0000	0.00%	-1.09%	5.38%	Arithmetic Average Return			**8.54%**
15	4-Oct-04	25.04		-0.0091	19.69%	0.9909	-0.91%	-0.46%	-2.99%	Geometric Average Return			**6.16%**
16	27-Sep-04	25.27		0.0346	19.68%	1.0352	3.52%	-0.13%	-1.45%				
17	20-Sep-04	24.41		-0.0082	19.62%	0.9919	-0.81%	-0.50%	-5.50%				
18	13-Sep-04	24.61		0.0008	20.52%	1.0008	0.08%	-5.59%	-1.57%				
19	7-Sep-04	24.59		0.0139	21.30%	1.0140	1.40%	0.05%	-7.74%				
20	30-Aug-04	24.25		-0.0127	21.25%	0.9874	-1.26%	-1.51%	-3.56%				
21	23-Aug-04	24.56		0.0123	22.29%	1.0124	1.24%	6.77%	-2.45%				
22	16-Aug-04	24.26		0.0066	22.29%	1.0066	0.66%	6.70%	3.10%				
23	9-Aug-04	24.10		-0.0041	22.42%	0.9959	-0.41%	8.68%	3.59%				
24	2-Aug-04	24.20		-0.0488	22.42%	0.9524	-4.76%	8.92%	6.62%				
25	26-Jul-04	25.41		0.0163	21.97%	1.0164	1.64%	11.44%	11.33%				
26	19-Jul-04	25.00		0.0198	22.11%	1.0200	2.00%	7.12%	7.43%				
27	12-Jul-04	24.51		-0.0138	22.02%	0.9863	-1.37%	5.12%	2.73%				
28	6-Jul-04	24.85		-0.0250	22.04%	0.9753	-2.47%	4.96%	4.11%				
29	28-Jun-04	25.48		0.0000	22.07%	1.0000	0.00%	10.49%	5.07%				
30	21-Jun-04	25.48		0.0079	22.30%	1.0079	0.79%	13.88%	8.10%				

31	14-Jun-04	25.28	0.0574	22.48%	1.0591	5.91%	10.44%	10.64%
32	7-Jun-04	23.87	0.0311	22.71%	1.0315	3.15%	11.34%	2.20%
33	1-Jun-04	23.14	-0.0107	22.86%	0.9893	-1.07%	12.33%	5.69%
34	24-May-04	23.39	0.0129	23.19%	1.0130	1.30%	9.56%	10.84%
35	17-May-04	23.09	0.0013	23.21%	1.0013	0.13%	9.89%	5.61%
36	10-May-04	23.06	0.0030	23.87%	1.0030	0.30%	4.46%	7.10%
37	3-May-04	22.99	-0.0134	24.07%	0.9867	-1.33%	1.17%	1.36%
38	26-Apr-04	23.30	-0.0527	24.05%	0.9487	-5.13%	3.49%	-0.34%
39	19-Apr-04	24.56	0.0903	23.67%	1.0945	9.45%	12.11%	5.92%
40	12-Apr-04	22.44	-0.0124	21.92%	0.9877	-1.23%	1.55%	0.31%
41	5-Apr-04	22.72	-0.0144	22.50%	0.9857	-1.43%	8.19%	0.44%
42	29-Mar-04	23.05	0.0322	22.76%	1.0327	3.27%	6.07%	7.15%
43	22-Mar-04	22.32	0.0158	22.59%	1.0159	1.59%	4.49%	0.31%
44	15-Mar-04	21.97	-0.0296	23.72%	0.9708	-2.92%	-4.23%	0.41%

FIGURE 11.6 Computing annualized return and risk.

moving average or *AVERAGE(H3:H54)*. Similarly, the geometric average return is the average of the 52-week moving window of the geometric returns, that is, cell M15 is simply *AVERAGE(I3:I54)*, where in cell I3, we have *(POWER(B3/B54,1/52)–1)*52*, the geometric average computation. The arithmetic growth rate is typically higher than the geometric growth rate when the returns period to period are volatile. Typically, the geometric growth rate (with a moving average window) should be used.

QUESTION

1. Compare and contrast between a discrete versus continuous decision variable when used in an optimization under uncertainty.

EXERCISE

1. Create an Excel model for a continuous optimization problem with the following parameters:
   a. A stock portfolio consisting of four individual stocks, each with its own return and risk profile—each return and risk value has its own distributional assumption that is correlated to one another.
   b. The optimization problem is to efficiently allocate your resources to those individual stocks such that the best bang for the buck is achieved—use a Sharpe ratio (portfolio returns to risk ratio).
   c. Optimize this portfolio of stocks through the Sharpe ratio and progressively create and show the Markowitz efficient frontier of stock allocations.

PART Seven

Risk Mitigation

CHAPTER 12

What Is So Real About Real Options, and Why Are They Optional?

This chapter provides the reader a cursory look at and quick introduction to real options analysis. It explains why only running simulations, forecasting, and optimization are not sufficient in a comprehensive risk management paradigm; that is, time-series forecasting and Monte Carlo simulation are used for *identifying*, *predicting*, and *quantifying* risks. The question that should be asked is, so what and what next? Quantifying and understanding risk is one thing, but turning this information into *actionable intelligence* is another. Real options analysis, when applied appropriately, allows you to *value* risk, creating strategies to *mitigate* risk, and how to position yourself to *take advantage* of risk. It is highly recommended that you refer to *Real Options Analysis: Tools and Techniques, Second Edition* (Wiley Finance, 2005) also by the author, in order to learn more about the theoretical as well as pragmatic step-by-step computational details of real options analysis.

WHAT ARE REAL OPTIONS?

In the past, corporate investment decisions were cut and dried. Buy a new machine that is more efficient, make more products costing a certain amount, and if the benefits outweigh the costs, execute the investment. Hire a larger pool of sales associates, expand the current geographical area, and if the marginal increase in forecast sales revenues exceeds the additional salary and implementation costs, start hiring. Need a new manufacturing plant? Show that the construction costs can be recouped quickly and easily by the increase in revenues the plant will generate through new and improved products, and the initiative is approved.

However, real-life business conditions are a lot more complicated. Your firm decides to go with an e-commerce strategy, but multiple strategic paths exist. Which path do you choose? What are the options you have? If you choose the wrong path, how do you get back on the right track? How do you value and prioritize the paths that exist? You are a venture capitalist firm with multiple business plans to consider. How do you value a start-up firm with no proven track record? How do you structure a mutually beneficial investment deal? What is the optimal timing to a second or third round of financing?

> Business conditions are fraught with uncertainty and risks. These uncertainties hold with them valuable information. When uncertainty becomes resolved through the passage of time, managers can make the appropriate midcourse corrections through a change in business decisions and strategies. Real options incorporate this learning model, akin to having a strategic road map, whereas traditional analyses that neglect this managerial flexibility will grossly undervalue certain projects and strategies.

Real options are useful not only in valuing a firm through its strategic business options, but also as a strategic business tool in capital investment decisions. For instance, should a firm invest millions in a new e-commerce initiative? How does a firm choose among several seemingly cashless, costly, and unprofitable information-technology infrastructure projects? Should a firm indulge its billions in a risky research and development initiative? The consequences of a wrong decision can be disastrous or even terminal for certain firms. In a traditional discounted cash-flow model, these questions cannot be answered with any certainty. In fact, some of the answers generated through the use of the traditional discounted cash-flow model are flawed because the model assumes a static, one-time decision-making process whereas the real options approach takes into consideration the strategic managerial options certain projects create under uncertainty and management's flexibility in exercising or abandoning these options at different points in time, when the level of uncertainty has decreased or has become known over time.

The real options approach incorporates a learning model, such that management makes better and more informed strategic decisions when some levels of uncertainty are resolved through the passage of time. The discounted cash-flow analysis assumes a static investment decision and assumes that strategic decisions are made initially with no recourse to choose other pathways or options in the future. To create a good analogy of real options, visualize it as a strategic road map of long and winding roads with multiple perilous turns and branches along the way. Imagine the intrinsic and extrinsic value of having such a road map or global positioning system when navigating through unfamiliar territory, as well as having road signs at every turn to guide you in making the best and most informed driving decisions. Such a strategic map is the essence of real options.

The answer to evaluating such projects lies in real options analysis, which can be used in a variety of settings, including pharmaceutical drug development, oil and gas exploration and production, manufacturing, start-up valuation, venture capital investment, information technology infrastructure, research and development, mergers and acquisitions, e-commerce and e-business, intellectual capital development, technology development, facility expansion, business project prioritization, enterprise-wide risk management, business unit capital budgeting, licenses, contracts, intangible asset valuation, and the like. The following section illustrates some business cases and how real options can assist in identifying and capturing additional strategic value for a firm.

THE REAL OPTIONS SOLUTION IN A NUTSHELL

Simply defined, real options methodology is a systematic approach and integrated solution using financial theory, economic analysis, management science, decision sciences, statistics, and econometric modeling in applying options theory in valuing real physical assets, as opposed to financial assets, in a dynamic and uncertain business environment where business decisions are flexible in the context of strategic capital investment decision making, valuing investment opportunities, and project capital expenditures.

Real options are crucial in:

- Identifying different corporate investment decision pathways or projects that management can navigate given highly uncertain business conditions.
- Valuing each of the strategic decision pathways and what it represents in terms of financial viability and feasibility.
- Prioritizing these pathways or projects based on a series of qualitative and quantitative metrics.
- Optimizing the value of strategic investment decisions by evaluating different decision paths under certain conditions or using a different sequence of pathways that can lead to the optimal strategy.
- Timing the effective execution of investments and finding the optimal trigger values and cost or revenue drivers.
- Managing existing or developing new optionalities and strategic decision pathways for future opportunities.

ISSUES TO CONSIDER

Strategic options do have significant intrinsic value, but this value is only realized when management decides to execute the strategies. Real options theory assumes that management is logical and competent and that management acts in the best interests of the company and its shareholders through the maximization of wealth and minimization of risk of losses. For example, suppose a firm owns the rights to a piece of land that fluctuates dramatically in price. An analyst calculates the volatility of prices and recommends that management retain ownership for a specified time period, where within this period there is a good chance that the price of real estate will triple. Therefore, management owns a call option, an *option to wait* and defer sale for a particular time period. The value of the real estate is therefore higher than the value that is based on today's sale price. The difference is simply this option to wait. However, the value of the real estate will not command the higher value if prices do triple but management decides not to execute the option to sell. In that case, the price of real estate goes back to its original levels after the specified period, and then management finally relinquishes its rights.

> Strategic optionality value can only be obtained if the option is executed; otherwise, all the options in the world are worthless.

Was the analyst right or wrong? What was the true value of the piece of land? Should it have been valued at its explicit value on a deterministic case where you know what the price of land is right now, and therefore this is its value; or should it include some types of optionality where there is a good probability that the price of land could triple in value, hence, the piece of land is truly worth more than it is now and should therefore be valued accordingly? The latter is the real options view. The additional strategic optionality value can only be obtained if the option is executed; otherwise, all the options in the world are worthless. This idea of *explicit* versus *implicit* value becomes highly significant when management's compensation is tied directly to the actual performance of particular projects or strategies.

To further illustrate this point, suppose the price of the land in the market is currently \$10 million. Further, suppose that the market is highly liquid and volatile and that the firm can easily sell off the land at a moment's notice within the next 5 years, the same amount of time the firm owns the rights to the land. If there is a 50 percent chance the price will increase to \$15 million and a 50 percent chance it will decrease to \$5 million within this time period, is the property worth an expected value of \$10 million? If the price rises to \$15 million, management should be competent and rational enough to execute the option and sell that piece of land immediately to capture the additional \$5 million premium. However, if management acts inappropriately or decides to hold off selling in the hopes that prices will rise even further, the property value may eventually drop back down to \$5 million. Now, how much is this property really worth? What if there happens to be an *abandonment option*? Suppose there is a perfect counterparty to this transaction who decides to enter into a contractual agreement whereby, for a contractual fee, the counterparty agrees to purchase the property for \$10 million within the next 5 years, regardless of the market price and executable at the whim of the firm that owns the property. Effectively, a safety net has been created whereby the minimum floor value of the property has been set at \$10 million (less the fee paid); that is, there is a limited downside but an unlimited upside, as the firm can always sell the property at market price if it exceeds the floor value. Hence, this strategic *abandonment option* has increased the value of the property significantly. Logically, with this abandonment option in place, the value of the land with the option is definitely worth more than \$10 million. The real options approach seeks to value this additional inherent flexibility. Real options analysis allows the firm to determine how much this safety downside insurance or abandonment option is worth (i.e., what is the fair-market value of the contractual fee to obtain the option?), the optimal trigger price (i.e., at what price will it be optimal to sell the land?), and the optimal timing (i.e., what is the optimal amount of time to hold on to the land?).

IMPLEMENTING REAL OPTIONS ANALYSIS

First, it is vital to understand that real options analysis is *not* a simple set of equations or models. It is an *entire decision-making process* that enhances the traditional decision analysis approaches. It takes what has been tried-and-true financial analytics and evolves it to the next step by pushing the envelope of analytical techniques. In addition, it is vital to understand that 50 percent of the value in real options analysis is simply thinking about it. Another 25 percent of the value comes from the number crunching activities, while the final 25 percent comes from the results interpretation and explanation to management. Several issues should be considered when attempting to implement real options analysis:

- *Tools*. The correct tools are important. These tools must be more comprehensive than initially required because analysts will grow into them over time. Do not be restrictive in choosing the relevant tools. Always provide room for expansion. Advanced tools will relieve the analyst of detailed model building and let him or her focus instead on 75 percent of the value—thinking about the problem and interpreting the results. Chapter 13 illustrates the use of Real Options Super Lattice Solver (SLS) software and how even complex and customized real options problems can be solved with great ease.

- *Resources.* The best tools in the world are useless without the relevant human resources to back them up. Tools do not eliminate the analyst, but enhance the analyst's ability to effectively and efficiently execute the analysis. The right people with the right tools will go a long way. Because there are only a few true real options experts in the world who truly understand the theoretical underpinnings of the models as well as the practical applications, care should be taken in choosing the correct team. A team of real options experts is vital in the success of the initiative. A company should consider building a team of in-house experts to implement real options analysis and to maintain the ability for continuity, training, and knowledge transfer over time. Knowledge and experience in the theories, implementation, training, and consulting are the core requirements of this team of individuals. This is why training is vital. For instance, the Certified Risk Analyst certification program provides analysts and managers the opportunity to immerse themselves in the theoretical and real-life applications of simulation, forecasting, optimization, and real options (see www.realoptionsvaluation.com for more details).
- *Senior Management Buy-In.* The analysis buy-in has to be top-down where senior management drives the real options analysis initiative. A bottom-up approach where a few inexperienced junior analysts try to impress the powers that be will fail miserably.

INDUSTRY LEADERS EMBRACING REAL OPTIONS

Industries using real options as a tool for strategic decision making started with oil and gas and mining companies and later expanded into utilities, biotechnology, pharmaceuticals, and now into telecommunications, high-tech, and across all industries. The following examples relate how real options have been or should be used in different companies.

Automobile and Manufacturing Industry

In automobile and manufacturing, General Motors (GM) applies real options to create *switching options* in producing its new series of autos. This option is essentially to use a cheaper resource over a given period of time. GM holds excess raw materials and has multiple global vendors for similar materials with excess contractual obligations above what it projects as necessary. The excess contractual cost is outweighed by the significant savings of switching vendors when a certain raw material becomes too expensive in a particular region of the world. By spending the additional money in contracting with vendors and meeting their minimum purchase requirements, GM has essentially paid the premium on purchasing a switching option, which is important especially when the price of raw materials fluctuates significantly in different regions around the world. Having an option here provides the holder a hedging vehicle against pricing risks.

Computer Industry

In the computer industry, HP-Compaq used to forecast sales in foreign countries months in advance. It then configured, assembled, and shipped the highly specific configuration printers to these countries. However, given that demand changes rapidly and forecast figures are seldom correct, the preconfigured printers usually suffer the higher inventory holding cost or the cost of technological obsolescence. HP-Compaq can create an *option to wait* and defer making any decisions too early through building assembly plants in these foreign countries. Parts can then be shipped and assembled in specific configurations when demand is known, possibly weeks in advance rather than months in advance. These parts can be shipped anywhere in the world and assembled in any configuration necessary, while excess parts are interchangeable across different countries. The premium paid on this option is building the assembly plants, and the upside potential is the savings in making wrong demand forecasts.

Airline Industry

In the airline industry, Boeing spends billions of dollars and several years to decide if a certain aircraft model should even be built. Should the wrong model be tested in this elaborate strategy, Boeing's competitors may gain a competitive advantage relatively quickly. Because so many technical, engineering, market, and financial uncertainties are involved in the decision-making process, Boeing can conceivably create an *option to choose* through parallel development of multiple plane designs simultaneously, knowing very well the increasing cost of developing multiple designs simultaneously with the sole purpose of eliminating all but one in the near future. The added cost is the premium paid on the option. However, Boeing will be able to decide which model to abandon or continue when these uncertainties and risks become known over time. Eventually, all the models will be eliminated save one. This way, the company can hedge itself against making the wrong initial decision and benefit from the knowledge gained through parallel development initiatives.

Oil and Gas Industry

In the oil and gas industry, companies spend millions of dollars to refurbish their refineries and add new technology to create an *option to switch* their mix of outputs among heating oil, diesel, and other petrochemicals as a final product, using real options as a means of making capital and investment decisions. This option allows the refinery to switch its final output to one that is more profitable based on prevailing market prices, to capture the demand and price cyclicality in the market.

Telecommunications Industry

In the telecommunications industry, in the past, companies like Sprint and AT&T installed more fiber-optic cable and other telecommunications infrastructure than any other company in order to create a *growth option* in the future by providing a secure and extensive network and to create a high barrier to entry, providing a first-to-market advantage. Imagine having to justify to the board of directors the need to spend billions of dollars on infrastructure that will not be used for years to come. Without the use of real options, this decision would have been impossible to justify.

Utilities Industry

In the utilities industry, firms have created an *option to execute* and an *option to switch* by installing cheap-to-build inefficient energy generator *peaker* plants to be used only when electricity prices are high and to shut down when prices are low. The price of electricity tends to remain constant until it hits a certain capacity utilization trigger level, when prices shoot up significantly. Although this occurs infrequently, the possibility still exists, and by having a cheap standby plant, the firm has created the option to turn on the switch whenever it becomes necessary, to capture this upside price fluctuation.

Real Estate Industry

In the real estate arena, leaving land undeveloped creates an option to develop at a later date at a more lucrative profit level. However, what is the optimal wait time or the optimal trigger price to maximize returns? In theory, one can wait for an infinite amount of time, and real options provide the solution for the optimal timing and optimal price trigger value.

Pharmaceutical Research and Development Industry

In pharmaceutical or research and development initiatives, real options can be used to justify the large investments in what seems to be cashless and unprofitable under the discounted cash-flow method but actually creates *compound expansion options* in the future. Under the myopic lenses of a traditional discounted cash-flow analysis, the high initial investment of, say, a billion dollars in research and development may return a highly uncertain projected few million dollars over the next few years. Management will conclude under a net present value analysis that the project is not financially feasible. However, a cursory look at the industry indicates that research and development is performed everywhere. Hence, management must see an intrinsic strategic value in research and development. How is this intrinsic strategic value quantified? A real options approach would optimally time and spread the billion dollar initial investment into a multiple-stage investment structure. At each stage, management has an *option to wait* and see what happens as well as the *option to abandon* or the *option to expand* into the subsequent stages. The ability to defer cost and proceed only if situations are permissible creates value for the investment.

High-Tech and e-Business Industry

In e-business strategies, real options can be used to prioritize different e-commerce initiatives and to justify those large initial investments that have an uncertain future. Real options can be used in e-commerce to create incremental investment stages compared to a large one-time investment (invest a little now, wait and see before investing more) as well as create *options to abandon* and other future growth options.

Mergers and Acquisition

In valuing a firm for acquisition, you should not only consider the revenues and cash flows generated from the firm's operations but also the strategic options that come with the firm. For instance, if the acquired firm does not operate up to expectations, an *abandonment option* can be executed where it can be sold for its intellectual property and other tangible assets. If the firm is highly successful, it can be spun off into other industries and verticals or new products and services can be eventually developed through the execution of an *expansion option*. In fact, in mergers and acquisition, several strategic options exist. For instance, a firm acquires other entities to enlarge its existing portfolio of products or geographic location or to obtain new technology (*expansion option*); or to divide the acquisition into many smaller pieces and sell them off as in the case of a corporate raider (*abandonment option*); or it merges to form a larger organization due to certain synergies and immediately lays off many of its employees (*contraction option*). If the seller does not value its real options, it may be leaving money on the negotiation table. If the buyer does not value these strategic options, it is undervaluing a potentially highly lucrative acquisition target.

All these cases where the high cost of implementation with no apparent payback in the near future seems foolish and incomprehensible in the traditional discounted cash-flow sense are fully justified in the real options sense when taking into account the strategic options the practice creates for the future, the uncertainty of the future operating environment, and management's flexibility in making the right choices at the appropriate time.

WHAT THE EXPERTS ARE SAYING

The trend in the market is quickly approaching the acceptance of real options, as can be seen from the following sample publication excerpts below.

According to an article in *Bloomberg Wealth Manager* (November 2001):

> Real options provide a powerful way of thinking and I can't think of any analytical framework that has been of more use to me in the past five years that I've been in this business.

According to a *Wall Street Journal* article (February 2000):

> Investors who, after its IPO in 1997, valued only Amazon.com's prospects as a book business would have concluded that the stock was significantly overpriced and missed the subsequent extraordinary price appreciation. Though assessing the value of real options is challenging, without doing it an investor has no basis for deciding whether the current stock price incorporates a reasonable premium for real options or whether the shares are simply overvalued.

CFO Europe (July/August 1999) cites the importance of real options in that:

> A lot of companies have been brainwashed into doing their valuations on a one-scenario discounted cash-flow basis and sometimes our recommendations are not what intuition would suggest, and that's where the real surprises come from—and with real options, you can tell exactly where they came from.

According to a *Business Week* article (June 1999):

> The real options revolution in decision making is the next big thing to sell to clients and has the potential to be the next major business breakthrough. Doing this analysis has provided a lot of intuition you didn't have in the past and that as it takes hold, it's clear that a new generation of business analysts will be schooled in options thinking. Silicon Valley

> is fast embracing the concepts of real options analytics, in its tradition of fail fast so that other options may be sought after.

In *Products Financiers* (April 1999):

> Real options is a new and advanced technique that handles uncertainty much better than traditional evaluation methods. Because many managers feel that uncertainty is the most serious issue they have to face, there is no doubt that this method will have a bright future as any industry faces uncertainty in its investment strategies.

A *Harvard Business Review* article (September/October 1998) hits home:

> Unfortunately, the financial tool most widely relied on to estimate the value of a strategy is the discounted cash flow, which assumes that we will follow a predetermined plan regardless of how events unfold. A better approach to valuation would incorporate both the uncertainty inherent in business and the active decision making required for a strategy to succeed. It would help executives to think strategically on their feet by capturing the value of doing just that—of managing actively rather than passively and real options can deliver that extra insight.

This chapter and the next provide a novel approach to applying real options to answering these issues and more. In particular, a real options framework is presented. It takes into account managerial flexibility in adapting to everchanging strategic, corporate, economic, and financial environments over time as well as the fact that in the real business world opportunities and uncertainty exist and are dynamic in nature. This book provides a real options process framework to identify, justify, time, prioritize, value, and manage corporate investment strategies under uncertainty in the context of applying real options.

The recommendations, strategies, and methodologies outlined here are not meant to replace traditional discounted cash-flow analysis but to complement it when the situation and the need arise. The entire analysis could be done, or parts of it could be adapted to a more traditional approach. In essence, the process methodology outlined starts with traditional analyses and continues with value- and insight-adding analytics, including Monte Carlo simulation, forecasting, real options analysis, and portfolio optimization. The real options approach outlined is not the only viable alternative nor will it provide a set of infallible results. However, if utilized correctly with the traditional approaches, it may lead to a set of more robust, accurate, insightful, and plausible results. The insights generated through real options analytics provide significant value in understanding a project's true strategic value.

CRITICISMS, CAVEATS, AND MISUNDERSTANDINGS IN REAL OPTIONS

Before embarking on a real options analysis, analysts should be aware of several caveats. The following five requirements need to be satisfied before a real options analysis can be run:

1. *A financial model must exist.* Real options analysis requires the use of an existing discounted cash-flow model, as real options build on the existing tried-and-true approaches of current financial modeling techniques. If a model does not exist, it means that strategic decisions have already been made and no financial justifications are required, and hence, there is no need for financial modeling or real options analysis.
2. *Uncertainties must exist.* Without uncertainty, the option value is worthless. If everything is known for certain in advance, then a discounted cash-flow model is sufficient. In fact, when volatility (a measure of risk and uncertainty) is zero, everything is certain, the real options value is zero, and the total strategic value of the project or asset reverts to the net present value in a discounted cash-flow model.
3. *Uncertainties must affect decisions when the firm is actively managing the project and these uncertainties must affect the results of the financial model.* These uncertainties will then become risks, and real options can be used to hedge the downside risk and take advantage of the upside uncertainties.
4. *Management must have strategic flexibility or options to make midcourse corrections when actively managing the projects.* Otherwise, do not apply real options analysis when there are no options or management flexibility to value.
5. *Management must be smart enough and credible enough to execute the options when it becomes optimal to do so.* Otherwise, all the options in the world are useless unless they are executed appropriately, at the right time, and under the right conditions.

There are also several criticisms against real options analysis. It is vital that the analyst understands what they are and what the appropriate responses are, prior to applying real options.

- *Real options analysis is merely an academic exercise and is not practical in actual business applications.* Nothing is further from the truth. Although it was true in the past that real options analysis was merely academic, many corporations have begun to embrace and apply real options analysis. Also, its concepts are very pragmatic, and with the use of the Real Options Super Lattice Solver software, even very difficult problems can be easily solved, as will become evident later in the next chapter. This book and software have helped bring the theoretical a lot closer to practice. Firms are using it and universities are teaching it. It is only a matter of time before real options analysis becomes part of normal financial analysis.

- *Real options analysis is just another way to bump up and incorrectly increase the value of a project to get it justified.* Again, nothing is further from the truth. If a project has significant strategic options but the analyst does not value them appropriately, he or she is leaving money on the table. In fact, the analyst will be incorrectly undervaluing the project or asset. Also, one of the foregoing requirements states that one should never run real options analysis unless strategic options and flexibility exist. If they do not exist, then the option value is zero, but if they do exist, neglecting their valuation will grossly and significantly underestimate the project or asset's value.
- *Real options analysis ends up choosing the highest risk projects, as the higher the volatility, the higher the option value.* This criticism is also incorrect. The option value is zero if no options exist. However, if a project is highly risky and has high volatility, then real options analysis becomes more important; that is, if a project is strategic but is risky, then you better incorporate, create, integrate, or obtain strategic real options to reduce and hedge the downside risk and take advantage of the upside uncertainties. Therefore, this argument is actually heading in the wrong direction. It is not that real options will overinflate a project's value, but for risky projects, you should create or obtain real options to reduce the risk and increase the upside, thereby increasing the total strategic value of the project. Also, although an option value is always greater than or equal to zero, sometimes the cost to obtain certain options may exceed its benefit, making the entire strategic value of the option negative, although the option value itself is always zero or positive.

So, it is incorrect to say that real options will always increase the value of a project or only risky projects are selected. People who make these criticisms do not truly understand how real options work. However, having said that, real options analysis is just another financial analysis tool, and the old axiom of *garbage in garbage out* still holds. But if care and due diligence are exercised, the analytical process and results can provide highly valuable insights. In fact, this author believes that 50 percent (rounded, of course) of the challenge and value of real options analysis is simply thinking about it. Understanding that you have options, or obtaining options to hedge the risks and take advantage of the upside, and to think in terms of strategic options, is half the battle. Another 25 percent of the value comes from actually running the analysis and obtaining the results. The final 25 percent of the value comes from being able to explain it to management, to your clients, and to yourself, such that the results become actionable, and not merely another set of numbers.

QUESTIONS

1. Create your own definition of real options analysis; that is, define real options analysis in a paragraph.
2. What are some of the possible approaches used to solve a real options analysis problem?
3. In choosing the right methodology to be used in a real options analysis, what are some of the key requirements that should be considered?
4. What are the necessary conditions that must exist before real options analysis can be applied on a project?
5. What is the major limitation of only using Monte Carlo simulation to perform risk analysis?

CHAPTER 13

The Black Box Made Transparent: Real Options Super Lattice Solver Software

Now that you are confident with the applicability of real options, it is time to move on and use the Real Options Super Lattice Solver (SLS) software in the enclosed CD-ROM. As shown in Chapter 12, applying real options is not an easy task. The use of software-based models allows the analyst to apply a consistent, well-tested, and replicable set of models. It reduces computational errors and allows the user to focus more on the process and problem at hand rather than on building potentially complex and mathematically intractable models. This chapter provides a good starting point with an introduction to the Super Lattice Solver software. For more details on using the software, consult the user manual, whereas for more technical, theoretical, and practical details of real options analysis, consult *Real Options Analysis: Tools and Techniques, Second Edition* (Wiley Finance, 2005). The materials covered in this chapter assume that the reader is sufficiently well versed in the basics of real options analytics.

The enclosed CD-ROM has a 30-day trial version of the Super Lattice Solver and Risk Simulator software. For professors, please contact the author for complimentary semester-long licenses for you and your students for installation in computer labs if this text and associated software are used in an entire class. The remainder of this chapter and relevant examples require the use of these software applications. To install the Super Lattice Solver software, insert the CD and wait for the setup program to start. If it does not start automatically, browse the content of the CD and double-click on the *CDAutorun.exe* file and follow the simple on-screen instructions. You must be connected to the Internet before you can download and install the latest version of the software. Click on Install the Super Lattice Solver software.

When prompted, enter the following user name and license key for a 30-day trial of the SLS software:

Name: **30 Day License** License Key: **513C-27D2-DC6B-9666**

Another license key is required to permanently unlock and use the software, and the license can be purchased by going to www.realoptionsvaluation.com.

After installing the software, verify that the installation was successful by clicking on and making sure that the following folder exists: *Start* | *Programs* | *Real Options Valuation* | *Real Options Super Lattice Solver*. Note that the SLS software will work on most international Windows operating systems but requires a quick change in settings by clicking on *Start* | *Control Panel* | *Regional and Language Options*. Select *English (United States)*. This is required because the numbering convention is different in foreign countries (e.g., one thousand dollars and fifty cents is written as 1,000.50 in the United States versus 1.000,50 in certain European countries).

INTRODUCTION TO THE REAL OPTIONS SUPER LATTICE SOLVER SOFTWARE

The Real Options Super Lattice Software comprises several modules, including the Single Super Lattice Solver (SLS), Multiple Super Lattice Solver (MSLS), Multinomial Lattice Solver (MNLS), SLS Excel Solution, and SLS Functions. These modules are highly powerful and customizable binomial and multinomial lattice solvers and can be used to solve many types of options (including the three main families of options: *real options*, which deals with physical and intangible assets; *financial options*, which deals with financial assets and the investments of such assets; and *employee stock options*, which deals with financial assets provided to employees within a corporation). This text illustrates some sample real options, financial options, and employee stock options applications that users will encounter most frequently. The following are the modules in the *Real Options Super Lattice Software*:

- The SLS is used primarily for solving options with a *single underlying asset* using binomial lattices. Even highly complex options with a single underlying asset can be solved using the SLS. The types of options solved include American, Bermudan, and European options to abandon, choose, contract, defer, execute, expand, and wait, as well as any customized combinations of these options with changing inputs over time.
- The MSLS is used for solving options with *multiple underlying assets* and sequential compound options with *multiple phases* using binomial lattices. Highly complex options with multiple underlying assets and phases can be solved using the MSLS. The types of options solved include multiple-phase stage-gate sequential compound options, simultaneous compound options, switching options, multiple-asset chooser options,

and customized combinations of phased options with all the option types solved using the SLS module previously described.
- The MNLS uses *multinomial lattices* (trinomial, quadranomial, pentanomial) to solve specific options that cannot be solved using binomial lattices. The options solved include mean-reverting, jump-diffusion, and rainbow two-asset options.
- The SLS Excel Solution implements the SLS and MSLS computations within the Excel environment, allowing users to access the SLS and MSLS functions directly in Excel. This feature facilitates model building, formula and value linking and embedding, and running simulations, and provides the user sample templates to create such models.
- The SLS Functions are additional real options and financial options models accessible directly through Excel. This module facilitates model building, linking and embedding, and running simulations.

The SLS software is created by the author and accompanies the materials presented at different training courses on real options, simulation, employee stock options valuation, and Certified Risk Analyst programs taught by the author. While the software and its models are based on his books, the training courses cover the real options subject matter in more depth, including the solution of sample business cases and the framing of real options of actual cases. It is highly suggested that the reader familiarizes him- or herself with the fundamental concepts of real options in Chapters 6 and 7 of *Real Options Analysis, Second Edition* (Wiley Finance, 2005) prior to attempting an in-depth real options analysis using the software. Note that the first edition of *Real Options Analysis: Tools and Techniques* published in 2002 shows the Real Options Analysis Toolkit software, an older precursor to the Super Lattice Solver, also created by Dr. Johnathan Mun. The Super Lattice Solver Version 1.1 supersedes the Real Options Analysis Toolkit by providing the following enhancements, and is introduced in this second edition:

- All inconsistencies, computation errors, and bugs fixed and verified.
- Allowance of changing input parameters over time (customized options).
- Allowance of changing volatilities over time.
- Incorporation of Bermudan (vesting and blackout periods) and customized options.
- Flexible modeling capabilities in creating or engineering your own customized options.
- General enhancements to accuracy, precision, and analytical prowess.

As the creator of both the Super Lattice Solver and Real Options Analysis Toolkit software, the author suggests that the reader focuses on using the Super Lattice Solver as it provides many powerful enhancements and analytical flexibility over its predecessor, the older, less powerful, and less flexible Real Options Analysis Toolkit software.

SINGLE ASSET SUPER LATTICE SOLVER

Figure 13.1 illustrates the SLS module. After installing the software, the user can access the SLS by clicking on *Start* | *Programs* | *Real Options Valuation* | *Real Options Super Lattice Solver* | *Single Super Lattice Solver*. The SLS has several sections: Option Type, Basic Inputs, Custom Equations, Custom Variables, Benchmark, Result, and Create Audit Worksheet.

SLS Examples

To help you get started, several simple examples are in order. A simple European call option is computed in this example using SLS. To follow along, start this example file by selecting *Start* | *Programs* | *Real Options Valuation*

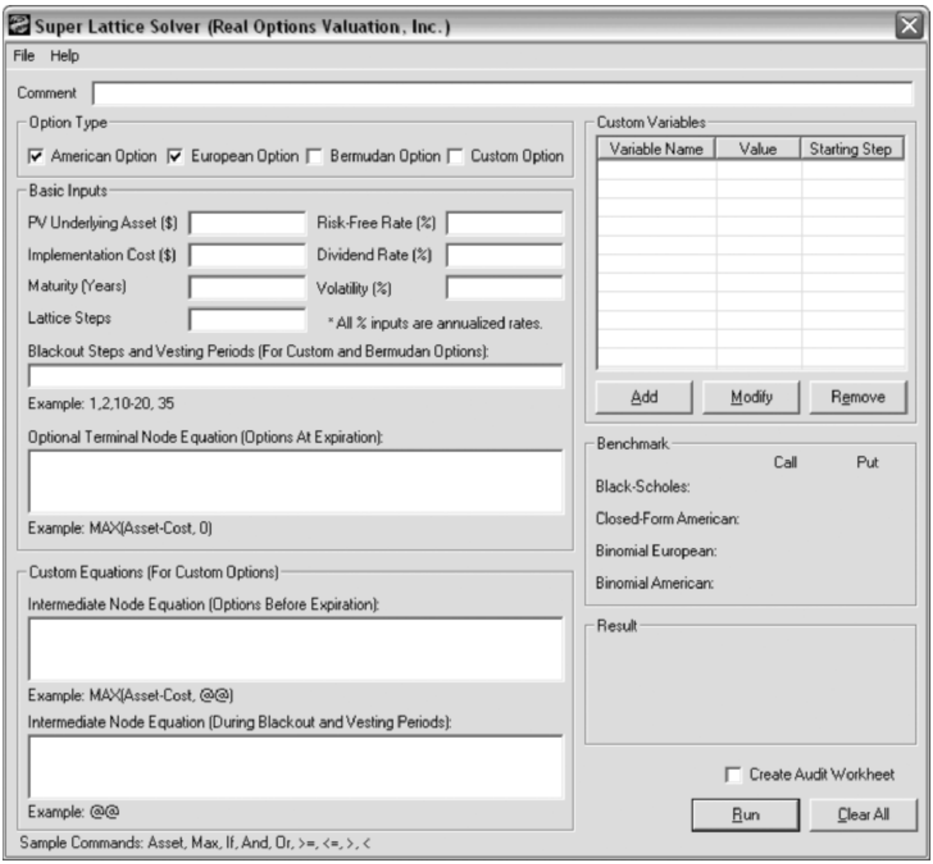


FIGURE 13.1 Single Super Lattice Solver (SLS).

| *Real Options Super Lattice Solver* | *Sample Files* | *Plain Vanilla Call Option I*. This example file will be loaded into the SLS software as seen in Figure 13.2. The starting PV Underlying Asset or starting stock price is \$100, and the Implementation Cost or strike price is \$100 with a 5-year maturity. The annualized risk-free rate of return is 5 percent, and the historical, comparable, or future expected annualized volatility is 10 percent. Click on RUN (or Alt-R) and a 100-step binomial lattice is computed and the results indicate a value of \$23.3975 for both the European and American call options. Benchmark values using Black–Scholes and Closed-Form American approximation models as well as standard plain-vanilla Binomial American and Binomial European Call and Put Options with 1,000-step binomial lattices are also computed. Notice that only the American and European options are

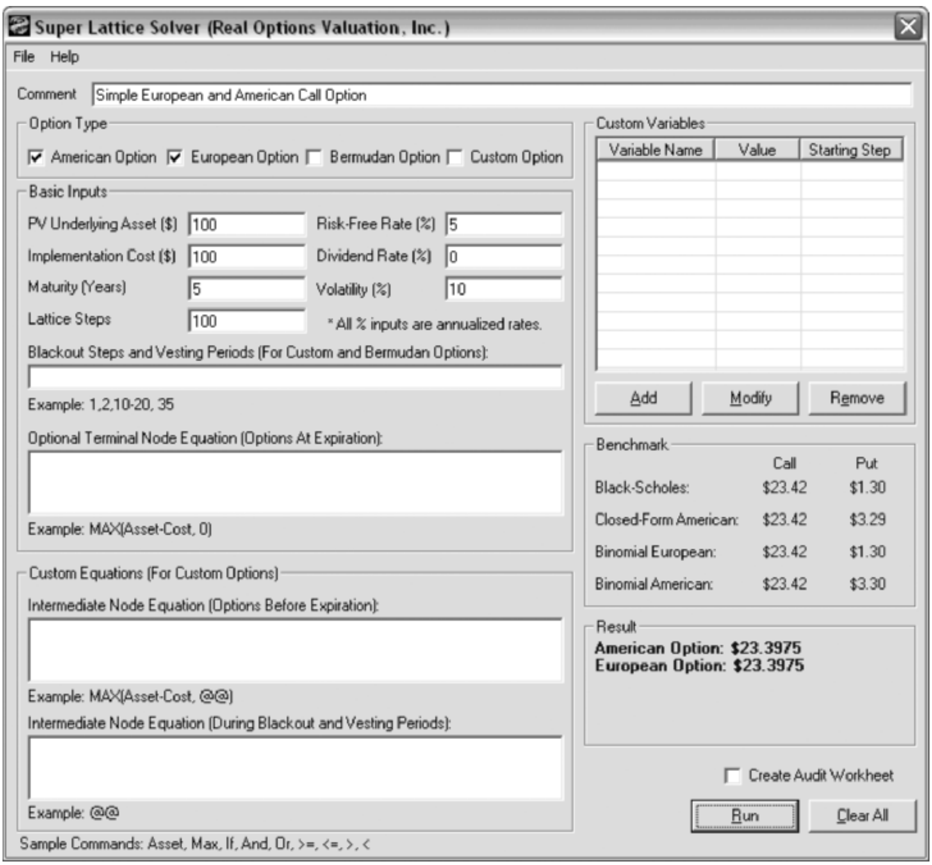


FIGURE 13.2 SLS results of a simple European and American Call Option.

selected and the computed results are for these simple plain-vanilla American and European call options.

The benchmark results use both closed-form models (Black–Scholes and Closed-Form Approximation models) and 1,000-step binomial lattices on plain-vanilla options. You can change the steps to *1000* in the basic inputs section to verify that the answers computed are equivalent to the benchmarks as seen in Figure 13.3. Notice that, of course, the values computed for the American and European options are identical to each other and identical to the benchmark values of $23.4187, as it is never optimal to exercise a standard plain-vanilla call option early if there are no dividends. Be aware that the higher the lattice steps, the longer it takes to compute the results. It is advisable to start with lower lattice steps to make sure the analysis is robust and then progressively increase lattice steps to check for results convergence.

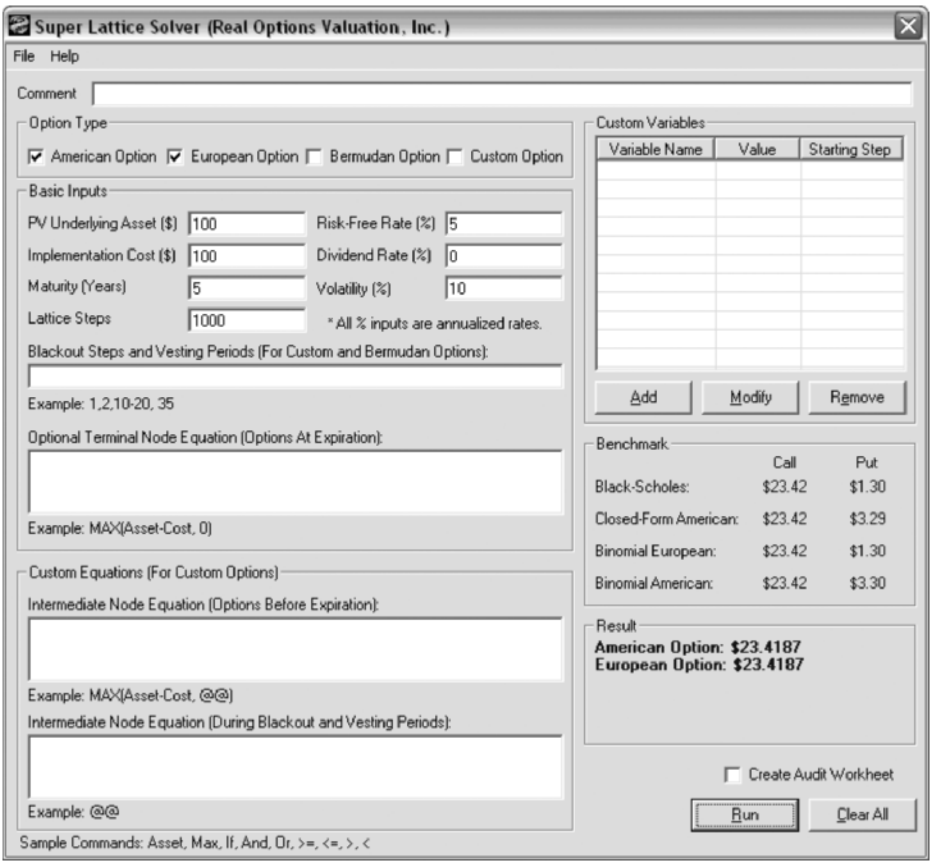


FIGURE 13.3 SLS comparing results with benchmarks.

Alternatively, you can enter Terminal and Intermediate Node Equations for a call option to obtain the same results. Notice that using 100 steps and creating your own Terminal Node Equation of *Max(Asset-Cost,0)* and Intermediate Node Equation of *Max(Asset-Cost,@@)* will yield the same answer. When entering your own equations, make sure that Custom Option is first checked.

> When entering your own equations, make sure that Custom Option is first checked.

Figure 13.4 illustrates how the analysis is done. The example file used in this illustration is: *Plain Vanilla Call Option III*. Notice that the value

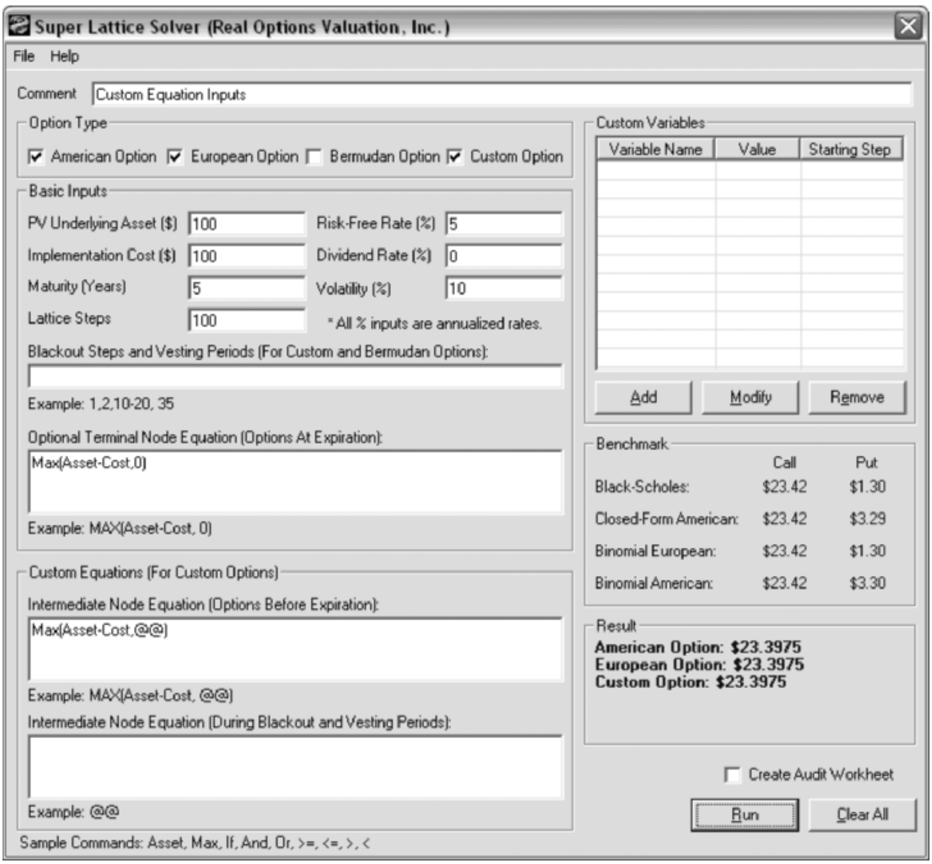


FIGURE 13.4 Custom equation inputs.

$23.3975 in Figure 13.4 agrees with the value in Figure 13.2. The Terminal Node Equation is the computation that occurs at maturity, while the Intermediate Node Equation is the computation that occurs at all periods prior to maturity, and is computed using backward induction. The symbol "@@" represents "keeping the option open," and is often used in the Intermediate Node Equation when analytically representing the fact that the option is not executed but kept open for possible future execution. Therefore, in Figure 13.4, the Intermediate Node Equation *Max(Asset-Cost,@@)* represents the profit maximization decision of either executing the option or leaving it open for possible future execution. In contrast, the Terminal Node Equation of *Max(Asset-Cost,0)* represents the profit maximization decision at maturity of either executing the option if it is in-the-money, or allowing it to expire worthless if it is at-the-money or out-of-the-money.

In addition, you can create an Audit Worksheet in Excel to view a sample 10-step binomial lattice by checking the box Create Audit Worksheet. For instance, loading the example file *Plain Vanilla Call Option I* and selecting the box creates a worksheet as seen in Figure 13.5. Several items on this audit worksheet are noteworthy:

- The audit worksheet generated will show the first 10 steps of the lattice, regardless of how many you enter; that is, if you enter 1,000 steps, the first 10 steps will be generated. If a complete lattice is required, simply enter 10 steps in the SLS and the full 10-step lattice will be generated instead. The Intermediate Computations and Results are for the Super Lattice, based on the number of lattice steps entered, and not based on the 10-step lattice generated. To obtain the Intermediate Computations for 10-step lattices, simply rerun the analysis inputting 10 as the lattice steps. This way, the audit worksheet generated will be for a 10-step lattice, and the results from SLS will now be comparable (Figure 13.6).
- The worksheet only provides values as it is assumed that the user was the one who entered in the terminal and intermediate node equations, hence there is really no need to re-create these equations in Excel again. The user can always reload the SLS file and view the equations or print out the form if required (by clicking on *File* | *Print*).

The software also allows you to save or open analysis files; that is, all the inputs in the software will be saved and can be retrieved for future use. The results will not be saved because you may accidentally delete or change an input and the results will no longer be valid. In addition, rerunning the super lattice computations will only take a few seconds, and it is advisable for you to always rerun the model when opening an old analysis file.

You may also enter in Blackout Steps. These are the steps on the super lattice that will have different behaviors than the terminal or intermediate

Option Valuation Audit Sheet

Assumptions		**Intermediate Computations**	
PV Asset Value ($)	$100.00	Stepping Time (dt)	0.0500
Implementation Cost ($)	$100.00	Up Step Size (up)	1.0226
Maturity (Years)	5.00	Down Step Size (down)	0.9779
Risk-free Rate (%)	5.00%	Risk-neutral Probability	0.5504
Dividends (%)	0.00%		
Volatility (%)	10.00%	**Results**	
Lattice Steps	100	Lattice Result	23.40
Option Type	European		

Terminal Equation	Max (Asset-Cost, 0)
Intermediate Equation	@@
Intermediate Equation (Blackouts)	@@

Underlying Asset Lattice

										125.06
									122.29	
								119.59		119.59
							116.94		116.94	
						114.36		114.36		114.36
					111.83		111.83		111.83	
				109.36		109.36		109.36		109.36
			106.94		106.94		106.94		106.94	
		104.57		104.57		104.57		104.57		104.57
	102.26		102.26		102.26		102.26		102.26	
100.00		100.00		100.00		100.00		100.00		100.00
	97.79		97.79		97.79		97.79		97.79	
		95.63		95.63		95.63		95.63		95.63
			93.51		93.51		93.51		93.51	
				91.44		91.44		91.44		91.44
					89.42		89.42		89.42	
						87.44		87.44		87.44
							85.51		85.51	
								83.62		83.62
									81.77	
										79.96

Option Valuation Lattice

										45.33
									42.81	
								40.36		39.96
							37.97		37.58	
						35.66		35.27		34.87
					33.43		33.04		32.64	
				31.27		30.88		30.49		30.09
			29.18		28.80		28.41		28.02	
		27.18		26.79		26.41		26.02		25.64
	25.25		24.87		24.49		24.11		23.73	
23.40		23.03		22.65		22.28		21.90		21.52
	21.26		20.90		20.53		20.16		19.79	
		19.22		18.86		18.50		18.14		17.77
			17.28		16.93		16.58		16.22	
				15.45		15.10		14.76		14.41
					13.71		13.38		13.05	
						12.09		11.77		11.45
							10.58		10.27	
								9.19		8.89
									7.91	
										6.74

FIGURE 13.5 SLS-generated audit worksheet.

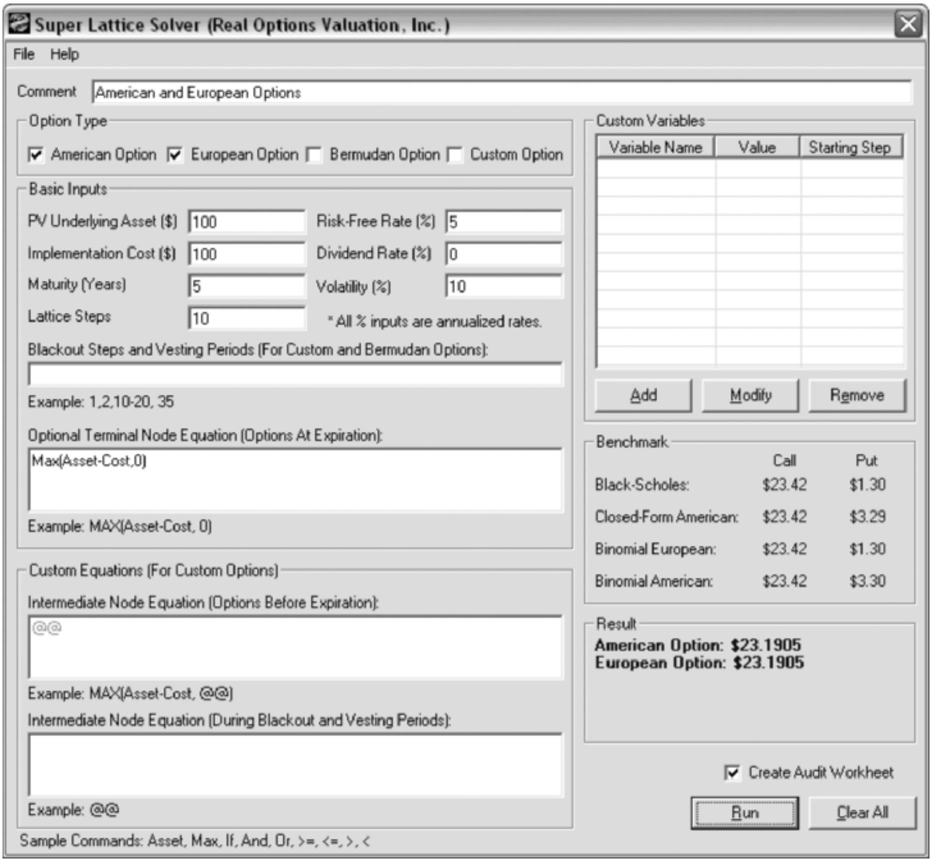


FIGURE 13.6 SLS results with a 10-step lattice.

steps. For instance, you can enter 1000 as the lattice steps, and enter 0-400 as the blackout steps, and some Blackout Equation (e.g., @@). This means that for the first 400 steps, the option holder can only keep the option open. Other examples include entering: 1, 3, 5, 10 if these are the lattice steps where blackout periods occur. You will have to calculate the relevant steps within the lattice where the blackout exists. For instance, if the blackout exists in years 1 and 3 on a 10-year, 10-step lattice, then steps 1, 3 will be the blackout dates. This blackout step feature comes in handy when analyzing options with holding periods, vesting periods, or periods where the option cannot be executed. Employee stock options have blackout and vesting periods, and certain contractual real options have periods during which the option cannot be executed (e.g., cooling-off periods, or proof of concept periods).

If equations are entered into the Terminal Node Equation box and American, European, or Bermudan Options are chosen, the Terminal Node Equation you entered will be the one used in the super lattice for the terminal nodes. However, for the intermediate nodes, the American option assumes the same Terminal Node Equation plus the ability to keep the option open; the European option assumes that the option can only be kept open and not executed; while the Bermudan option assumes that during the blackout lattice steps, the option will be kept open and cannot be executed. If you also enter the Intermediate Node Equation, the Custom Option should first be chosen (otherwise you cannot use the Intermediate Node Equation box). The Custom Option result uses all the equations you have entered in Terminal, Intermediate, and Intermediate during Blackout sections.

The Custom Variables list is where you can add, modify, or delete custom variables, the variables that are required beyond the basic inputs. For instance, when running an abandonment option, you need the salvage value. You can add this value in the Custom Variables list, provide it a name (a variable name must be a single word), the appropriate value, and the starting step when this value becomes effective. For example, if you have multiple salvage values (i.e., if salvage values change over time), you can enter the same variable name (e.g., salvage) several times, but each time, its value changes and you can specify when the appropriate salvage value becomes effective. For instance, in a 10-year, 100-step super lattice problem where there are two salvage values—\$100 occurring within the first 5 years and increases to \$150 at the beginning of Year 6—you can enter two salvage variables with the same name, \$100 with a starting step of 0, and \$150 with a starting step of 51. Be careful here as Year 6 starts at step 51 and not 61; that is, for a 10-year option with a 100-step lattice, we have: Steps 1–10 = Year 1; Steps 11–20 = Year 2; Steps 21–30 = Year 3; Steps 31–40 = Year 4; Steps 41–50 = Year 5; Steps 51–60 = Year 6; Steps 61–70 = Year 7; Steps 71–80 = Year 8; Steps 81–90 = Year 9; and Steps 91–100 = Year 10. Finally, incorporating 0 as a blackout step indicates that the option cannot be executed immediately.

MULTIPLE SUPER LATTICE SOLVER

The MSLS is an extension of the SLS in that the MSLS can be used to solve options with multiple underlying assets and multiple phases. The MSLS allows the user to enter multiple underlying assets as well as multiple valuation lattices (Figure 13.7). These valuation lattices can call to user-defined custom variables. Some examples of the types of options that the MSLS can be used to solve include:

- Sequential Compound Options (two-, three-, and multiple-phased sequential options).

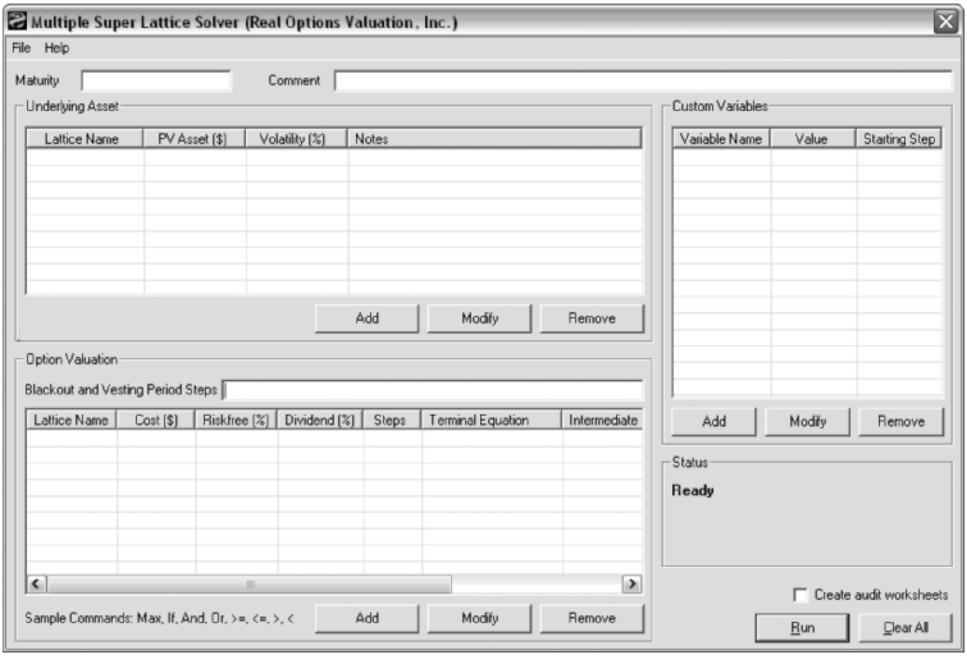


FIGURE 13.7 Multiple Super Lattice Solver.

- Simultaneous Compound Options (multiple assets with multiple simultaneous options).
- Chooser and Switching Options (choosing among several options and underlying assets).
- Floating Options (choosing between calls and puts).
- Multiple Asset Options (3D binomial option models).

The MSLS software has several areas including a *Maturity* and *Comment* area. The maturity value is a global value for the entire option, regardless of how many underlying or valuation lattices exist. The comment field is for your personal notes describing the model you are building. There is also a *Blackout and Vesting Period Steps* section and a *Custom Variables* list similar to the SLS. The MSLS also allows you to create audit worksheets.

To illustrate the power of the MSLS, a simple illustration is in order. Click on *Start* | *Programs* | *Real Options Valuation* | *Real Options Super Lattice Solver* | *Sample Files* | *MSLS—Two-Phased Sequential Compound Option*. Figure 13.8 shows the MSLS example loaded. In this simple example, a single underlying asset is created with two valuation phases.

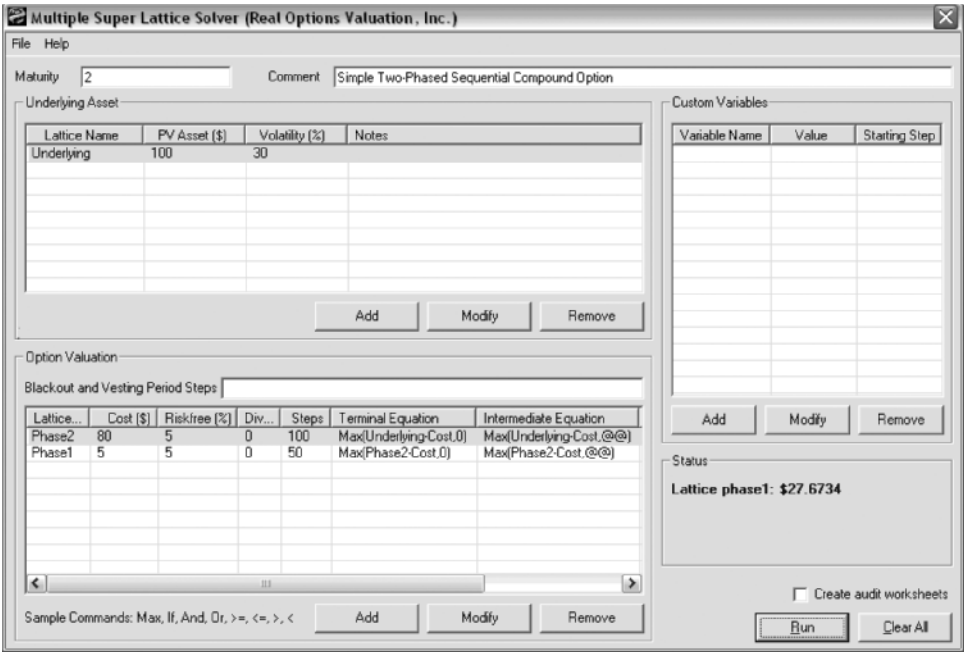


FIGURE 13.8 MSLS solution to a simple two-phased sequential compound option.

The strategy tree for this option is seen in Figure 13.9. The project is executed in two phases—the first phase within the first year costs \$5 million, while the second phase occurs within 2 years but only after the first phase is executed, and costs \$80 million, both in present value dollars. The PV Asset of the project is \$100 million (net present value is therefore \$15 million), and faces 30 percent volatility in its cash flows. The computed strategic value

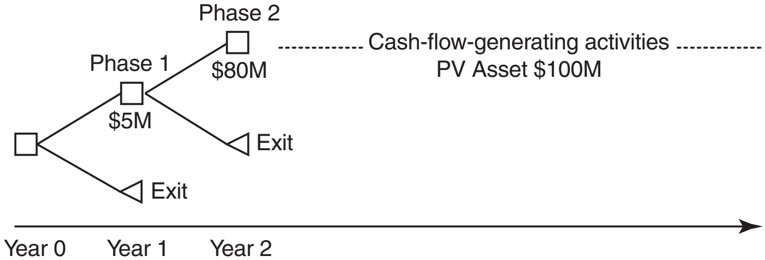


FIGURE 13.9 Strategy tree for two-phased sequential compound option.

using the MSLS is $27.67 million, indicating that there is a $12.67 million in option value; that is, spreading out and staging the investment into two phases has significant value (an expected value of $12.67 million to be exact).

MULTINOMIAL LATTICE SOLVER

The Multinomial Lattice Solver (MNLS) is another module of the Real Options Valuation's Super Lattice Solver software. The MNLS applies multinomial lattices—where multiple branches stem from each node—such as trinomials (three branches), quadranomials (four branches), and pentanomials (five branches). Figure 13.10 illustrates the MNLS module. The module has a Basic Inputs section, where all of the common inputs for the multinomials are listed. Then, there are four sections with four different multinomial applications complete with the additional required inputs and results for both American and European call and put options.

Figure 13.11 shows an example call and put option computation using trinomial lattices. To follow along, open the example file *MNLS—Simple Calls and Puts using Trinomial Lattices*. Note that the results shown in Figure 13.11 using a 50-step lattice are equivalent to the results shown in Figure 13.2 using a 100-step binomial lattice. In fact, a trinomial lattice or any other multinomial lattice provides identical answers to the binomial lattice at the limit, but convergence is achieved faster at lower steps. To illustrate,

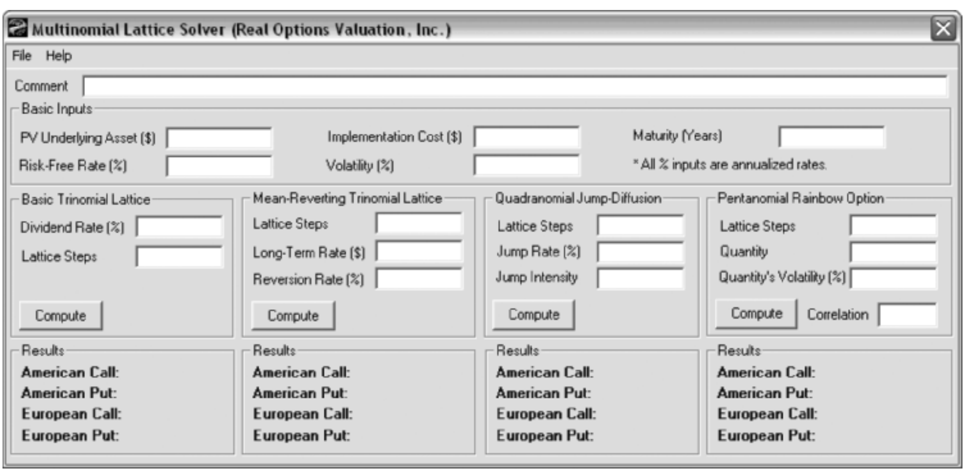


FIGURE 13.10 Multinomial lattice solver.

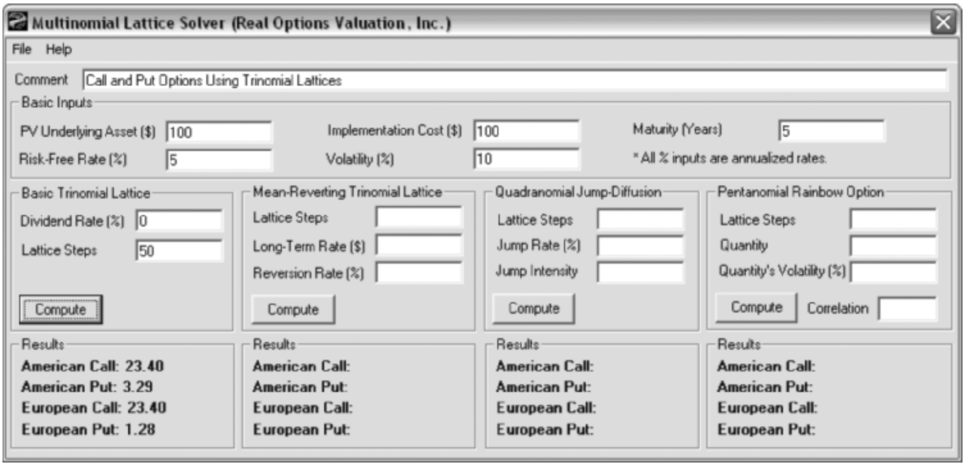


FIGURE 13.11 A simple call and put using trinomial lattices.

TABLE 13.1 Binomial Versus Trinomial Lattices

Steps	5	10	100	1,000	5,000
Binomial Lattice	\$30.73	\$29.22	\$29.72	\$29.77	\$29.78
Trinomial Lattice	\$29.22	\$29.50	\$29.75	\$29.78	\$29.78

Table 13.1 shows how the trinomial lattice of a certain set of input assumptions yields the correct option value with fewer steps than it takes for a binomial lattice. Because both yield identical results at the limit but trinomials are much more difficult to calculate and take a longer computation time, the binomial lattice is usually used instead. However, a trinomial is required only under one special circumstance: when the underlying asset follows a mean-reverting process.

With the same logic, quadranomials and pentanomials yield identical results as the binomial lattice with the exception that these multinomial lattices can be used to solve the following different special limiting conditions:

- *Trinomials*. Results are identical to binomials and are most appropriate when used to solve mean-reverting underlying assets.
- *Quadranomials*. Results are identical to binomials and are most appropriate when used to solve options whose underlying assets follow jump-diffusion processes.

- *Pentanomials*. Results are identical to binomials and are most appropriate when used to solve two underlying assets that are combined, called rainbow options (e.g., price and quantity are multiplied to obtain total revenues, but price and quantity each follows a different underlying lattice with its own volatility, but both underlying parameters could be correlated to one another).

SLS EXCEL SOLUTION

The SLS software also allows you to create your own SLS, MSLS, and Changing Volatility models in Excel using customized functions. This functionality is important because certain models may require linking from other spreadsheets or databases, run certain Excel macros and functions, or certain inputs need to be simulated, or inputs may change over the course of modeling your options. This Excel compatibility allows you the flexibility to innovate within the Excel spreadsheet environment. Specifically, the sample worksheet included in the software solves the SLS, MSLS, and Changing Volatility model.

To illustrate, Figure 13.12 shows a Customized Abandonment Option solved using SLS. The same problem can be solved using the *SLS Excel Solution* by clicking on *Start* | *Programs* | *Real Options Valuation* | *Real Options Super Lattice Solver* | *SLS Excel Solution*. The sample solution is seen in Figure 13.13. Notice the same results using the SLS versus the SLS Excel Solution file. You can use the template provided by simply clicking on *File* | *Save As* in Excel and use the new file for your own modeling needs.

Similarly, the MSLS can also be solved using the SLS Excel Solver. Figure 13.14 shows a complex multiple-phased sequential compound option solved using the SLS Excel Solver. The results shown here are identical to the results generated from the MSLS module (example file: *MSLS—Multiple Phased Complex Sequential Compound Option*). One small note of caution here is that if you add or reduce the number of option valuation lattices, make sure you change the function's link for the *MSLS Result* to incorporate the right number of rows; otherwise, the analysis will not compute properly. For example, the default shows three option valuation lattices, and by selecting the *MSLS Results* cell in the spreadsheet and clicking on *Insert* | *Function*, you will see that the function links to cells A24:H26 for these three rows for the *OVLattices* input in the function. If you add another option valuation lattice, change the link to A24:H27, and so forth. You can also leave the list of custom variables as is. The results will not be affected if these variables are not used in the custom equations.

Finally, Figure 13.15 shows a Changing Volatility and Changing Risk-free Rate Option. In this model, the volatility and risk-free yields are allowed

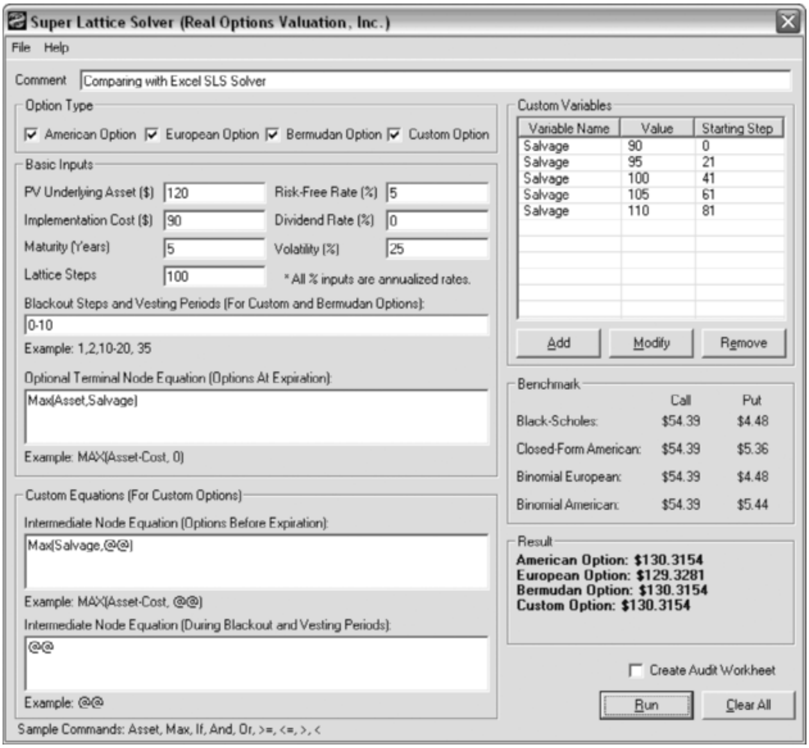


FIGURE 13.12 Customized abandonment option using SLS.

SUPER LATTICE SOLVER (SINGLE ASSET)

Option Type	0
PV Underlying Asset	$120.00
Annualized Volatility	25.00%
Maturity (Years)	5.00
Implementation Cost	$0.00
Risk-Free Rate	5.00%
Dividend Yield	0.00%
Lattice Steps	100
Terminal Equation	MAX (Asset, Salvage)
Intermediate Equation	MAX (Salvage, @@)
Intermediate Equation During Blackout	@@
Blackout Steps	0-10

Custom Variables List

Variable Name	Value	Starting Steps
Salvage	90.00	0
Salvage	95.00	21
Salvage	100.00	41
Salvage	105.00	61
Salvage	110.00	81

Super Lattice Solver Result **$130.3154**

Note: This is the Excel version of the Super Lattice Solver, useful when running simulations or when linking to and from other spreadsheets. Use this sample spreadsheet for your models. You can simply click on File, Save As to save as a different file and start using the model.
For the option type, set 0 = American, 1 = European, 2 = Bermudan, 3 = Custom
The function used is *SLSSingle*

FIGURE 13.13 Customized abandonment option using SLS Excel Solution.

MULTIPLE SUPER LATTICE SOLVER (MULTIPLE ASSET & MULTIPLE PHASES)

Maturity (Years)	5.00
Blackout Steps	0-20

MSLS Result $134.0802

Underlying Asset Lattices

Lattice Name	PV Asset	Volatility
Underlying	100.00	25.00

Custom Variables

Name	Value	Starting Steps
Salvage	100.00	31
Salvage	90.00	11
Salvage	80.00	0
Contract	0.90	0
Expansion	1.50	0
Savings	20.00	0

Option Valuation Lattices

Lattice Name	Cost	Riskfree	Dividend	Steps	Terminal Equation	Intermediate Equation	Intermediate Equation for Blackout
Phase3	50.00	5.00	0.00	50	Max(Underlying*Expansion-Cost,Underlying,Salvage)	Max(Underlying*Expansion-Cost,Salvage,@@)	@@
Phase2	0.00	5.00	0.00	30	Max(Phase3,Phase3*Contract+Savings,Salvage,0)	Max(Phase3*Contract+Savings,Salvage,@@)	@@
Phase1	0.00	5.00	0.00	10	Max(Phase2,Salvage,0)	Max(Salvage,@@)	@@

Note: This is the Excel version of the Multiple Super Lattice Solver, useful when running simulations or when linking to and from other spreadsheets.
Use this sample spreadsheet for your models. You can simply click on File, Save As to save as a different file and start using the model.
The function used is: *SLSMultiple*
One small note of caution here is that if you add or reduce the number of option valuation lattices make sure you change the function's link for the MSLS Result to incorporate the right number of rows otherwise the analysis will not compute properly. For example, the default shows 3 option valuation lattices and by selecting the MSLS Results cell F5 and clicking on Insert | Function, you will see that the function links to cells A24:H26 for these three rows for the OVLattices input in the function. If you add another option valuation lattice, change the link to A24:H27, and so forth.

FIGURE 13.14 Complex sequential compound option using SLS Excel Solver.

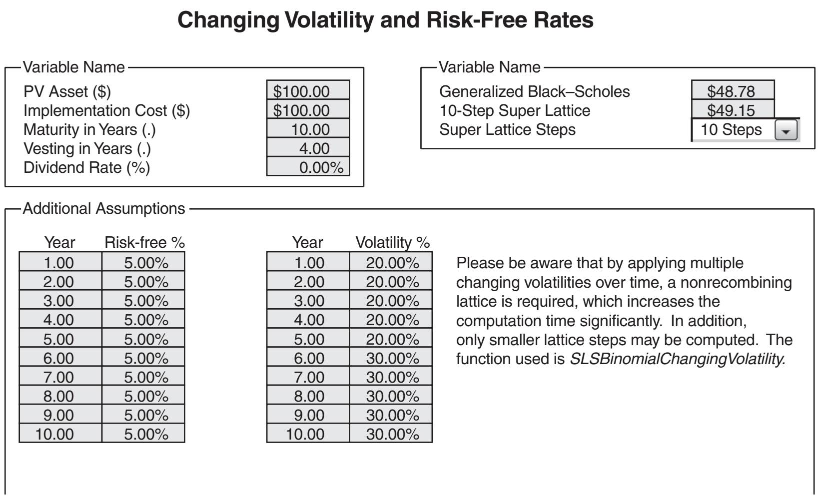


FIGURE 13.15 Changing volatility and risk-free rate option.

to change over time and a nonrecombining lattice is required to solve the option. In most cases, it is recommended that you create option models without the changing volatility term structure because getting a single volatility is difficult enough let alone a series of changing volatilities over time. If different volatilities that are uncertain need to be modeled, run a Monte Carlo simulation using the Risk Simulator software on volatilities instead. This model should only be used when the volatilities are modeled robustly and the volatilities are rather certain and change over time. The same advice applies to a changing risk-free rate term structure.

SLS FUNCTIONS

The software also provides a series of SLS functions that are directly accessible in Excel. To illustrate its use, start the SLS Functions by clicking on *Start* | *Programs* | *Real Options Valuation* | *Real Options Super Lattice Solver* | *SLS Functions*, and Excel will start. When in Excel, you can click on the function wizard icon or simply select an empty cell and click on *Insert* | *Function*. While in Excel's equation wizard, either select the *All* category or *Real Options Valuation*, the name of the company that developed the software. Here you will see a list of SLS functions (with SLS prefixes) that are ready for use in Excel. Figure 13.16 shows the Excel equation wizard.

Suppose you select the first function, *SLSBinomialAmericanCall* and hit OK. Figure 13.17 shows how the function can be linked to an existing Excel

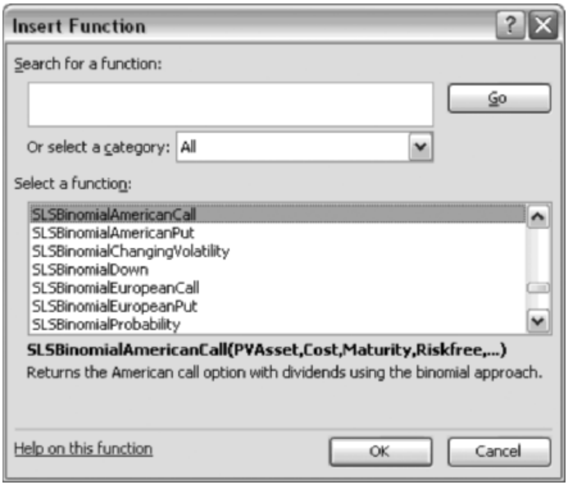


FIGURE 13.16 Excel's equation wizard.

model. The values in cells B1 to B7 can be linked from other models or spreadsheets, or can be created using Excel's Visual Basic for Applications (VBA) macros, or can be dynamic and changing as in when running a simulation. Another quick note of caution here is that certain SLS functions require many input variables, and Excel's equation wizard can only show five

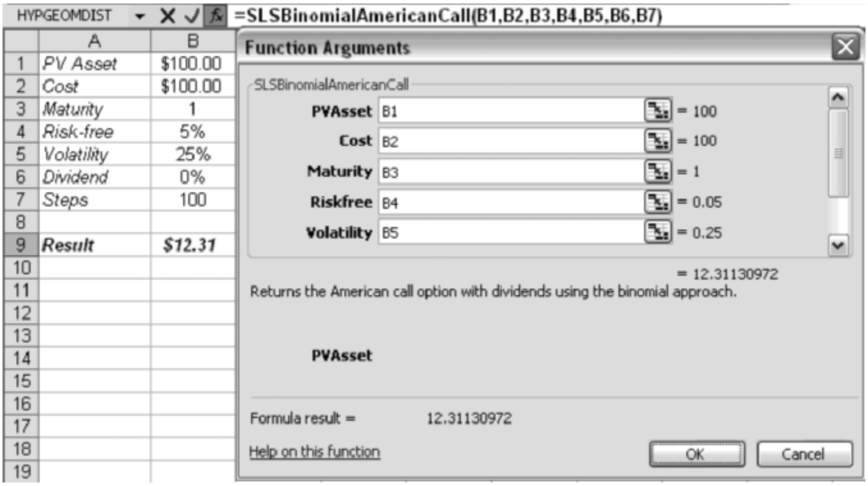


FIGURE 13.17 Using SLS functions in Excel.

variables at a time. Therefore, remember to scroll down the list of variables by clicking on the vertical scroll bar to access the rest of the variables.

LATTICE MAKER

Finally, the full version of the software comes with an advanced binomial *Lattice Maker* module. This Lattice Maker is capable of generating binomial lattices and decision lattices with visible formulas in an Excel spreadsheet. Figure 13.18 illustrates an example option generated using this module. The illustration shows the module inputs (you can obtain this module by clicking on *Start* | *Programs* | *Real Options Valuation* | *Real Options Super Lattice Solver* | *Lattice Maker*) and the resulting output lattice. Notice that the visible equations are linked to the existing spreadsheet, which means this module will come in handy when running Monte Carlo simulations or when used to link to and from other spreadsheet models. The results can also be used as a presentation and learning tool to peep inside the analytical black box of binomial lattices. Last but not least, a decision lattice with specific decision nodes indicating expected optimal times of execution of certain options are also available in this module. The results generated from this module are identical to those generated using the SLS and Excel functions, but with the added advantage of a visible lattice (lattices of up to 200 steps can be generated using this module). You are now equipped to start using the SLS software in building and solving real options, financial options, and employee stock options problems.

In conclusion, this chapter provides only a cursory look at how to solve real options using the SLS software. To truly understand, set up, and apply real options, review the author's *Real Options Analysis, Second Edition* (Wiley Finance, 2005).

Customized Real Option Results

ASSET(S)	VOLATILITY	RISK-FREE	DIVIDEND	MATURITY (Y)	STEPS
100.00	25.00%	5.00%	0.00%	1.00	5

UP SIZE	DOWN SIZE	UP PROB	DOWN PROB	DISC FACTOR	OPTION STYLE
1.118292981	0.894220045	0.516930442	0.483069558	0.990049834	AMERICAN

EXECUTION	EXPANSION	IMPLEMENTATION	CONTRACTION	SAVINGS	SALVAGE
0.00	1.30	25.00	0.90	25.00	120.00

100.00	111.83	125.06	139.85	156.39	174.90
	89.42	100.00	111.83	125.06	139.85
		79.96	89.42	100.00	111.83
			71.50	79.96	89.42
				63.94	71.50
					57.18

=MAX(A25*$B$23-$C$23,A25*$D$23+$E$23,$F$23,A25,(B32*$C$20+B34*$D$20)*$E$20)	132.62	143.31	158.51	178.56	202.36
	Continue	Continue	Continue	Continue	Continue
	120.82	123.94	130.03	140.34	156.81
	Continue	Continue	Continue	Continue	Continue
		120.00	120.00	121.70	125.65
		Abandon	Abandon	Continue	Contract
			120.00	120.00	120.00
			Abandon	Abandon	Abandon
				120.00	120.00
				Abandon	Abandon
					120.00
					Abandon

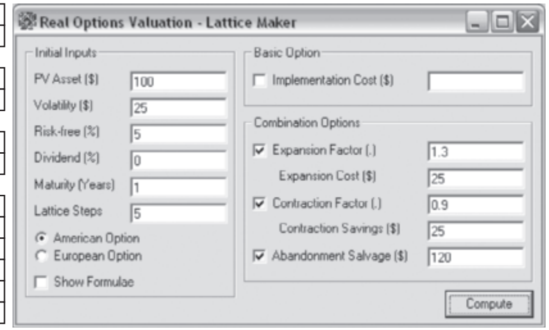


FIGURE 13.18 Lattice maker.

PART

Eight

More Industry Applications

CHAPTER 14

Extended Business Cases II: Real Estate, Banking, Military Strategy, Automotive Aftermarkets, Global Earth Observation Systems, and Employee Stock Options

This chapter provides six additional applied case studies. The first case is contributed by Robert Fourt and Professor Bill Rodney on real estate development using real options analysis techniques. For more details on the techniques used in this case—specifically relating to optimal timing, binomial lattices, and state-pricing approaches—refer to *Real Options Analysis, Second Edition* by Johnathan Mun (Wiley, 2005) and *Real Options Analysis Course* also by Johnathan Mun (Wiley, 2003). The second case is contributed by Professor Morton Glantz on the application of risk analysis, simulation, and optimization under uncertainty with respect to credit risk modeling in the banking industry. The third case is contributed by Dr. Tom Housel and Lt. Commander Cesar Rios on applying real options analysis and Monte Carlo simulation at the U.S. Navy and Department of Defense. The fourth case is on the billion-dollar automotive aftermarket by Larry Blair and Andy Roff. The fifth case is by Dr. Johnathan Mun and several senior engineers and analysts from Boeing, including Ken Cobleigh, on the Global Earth Observation System of Systems. Finally, the sixth case, also by Dr. Mun, looks at how Super Lattice Solver can be used to solve complex employee stock options and executive compensation with suboptimal exercise behaviors and performance barriers.

CASE STUDY: UNDERSTANDING RISK AND OPTIMAL TIMING IN A REAL ESTATE DEVELOPMENT USING REAL OPTIONS ANALYSIS

This case study is contributed by Robert Fourt (contact: Gerald Eve, 7 Vere Street, London W1G OJB, UK, +44(0)2074933338, rfourt@geraldeve.com*) and Bill Rodney (contact: Cass Business School, 106 Bunhill Row, London, EC1Y8TZ, UK, +44(0)2070408600,* whr@dial.pipex.com*). Robert is a partner within the planning and development team of UK-based real estate consultants, Gerald Eve. He specializes in development consultancy, providing advice on a wide range of schemes to corporate and public sector clients with a particular emphasis on strategy, finance, and project management. Gerald Eve is a multidisciplinary practice employing more than 300 people operating from a head office in central London and a regional network that spans the United Kingdom. The firm provides specialist advice in all real estate sectors. Bill is a senior lecturer in real estate finance at the Cass Business School, as well as undertaking research and providing advice to a number of institutions on real estate risk analysis, financing strategies, and the risk pricing of PPP/PFI projects. The Cass Business School (part of the City University) is a leading European center for finance research, investment management, and risk assessment and benefits from its location in the heart of London's financial district and involvement of leading practitioners in its teaching and research.*

Consideration of risk and its management is key in most real estate investment and development opportunities. Recognition of this, particularly in recent years, has led to various financial techniques being employed, including simulation analysis and Value at Risk (VaR), to assess various proposed transactions. The U.K. Investment Property Forum has sought to establish a real estate sector standard for risk. This standard for risk has provided a greater insight into the risk structure and returns on investments for management to review. Notwithstanding these approaches, they have nevertheless largely relied on traditional deterministic appraisals as a basis for assessing risk and return.

An addition to understanding the risks and returns of a project is to apply a real options analysis (ROA). In commercial real estate, the application of an ROA to date has largely been academically driven. While this has provided a strong theoretical base with complex numerical and analytical techniques employed, there has been limited practical application. This lack in some respects is surprising, given that real estate contains a multiplicity of embedded real options due to its intrinsic nature and that the sector operates under conditions of uncertainty. In particular, real estate development provides flexibility in deferring, commencing, or abandoning a project, which in turn are options that convey value.

This case example, which focuses on a large site in the town center of Croydon, 20 minutes from central London in the United Kingdom, highlights the differences of an investment's risk structure and average return when comparing a static net present value (NPV) to an ROA approach. It also illustrates the apparent irrationality of why land is left undeveloped in downtown locations despite the apparent redevelopment potential, an issue that has been the subject of several seminal real option real estate papers (see Notes at the end of this case).

The ROA approach for this example initially formed the basis for advice to the Council (local authority), which was working closely with an investor developer. For this case study, the analysis is from the perspective of the investor in seeking to understand the optimal timing for development and its associated risk structure. In order to maintain confidentiality and simplify certain steps, prices and issues referred to have been adapted.

The right or flexibility to develop (i.e., construct) land is a real option and this often comes in the form of an American call option. This case study utilizes a binomial lattice approach and methodology. The call option is combined with an American put to sell the site either to the Council at open market value (OMV) or as a result of compulsory purchase order (CPO). Therefore, the strategic decision is whether to defer, sell (i.e., abandon), or develop. This flexibility conveys value, which is not captured by a conventional deterministic or NPV appraisal.

A five-step ROA approach was adopted and comprised:

Stage I	Mapping or framing the problem.
Stage II	Base scoping appraisal (deterministic).
Stage III	Internal and external uncertainty inputs.
Stage IV	Real options quantitative analysis.
Stage V	Explanation and strategic decisions.

Three quantitative variations using a lattice approach were considered: a binomial lattice; state pricing; and a binomial lattice with two volatility variables. The reasoning for this approach is explained later. A Monte Carlo analysis was undertaken at both the deterministic analysis (Stage II) and with the ROA (Stage IV), which further illustrates the risk profile comparison between real options and NPV.

The lattice approach allows for decisions to be taken at each node. This features provides an investor with the ability to determine the optimal timing with respect to development, or to defer, or to abandon (disposal of the property).

The basic simplified details of this case study are as follows:

- An undeveloped town center site of approximately 2.43 ha (6 acres) adjacent to a major public transport interchange.

- A comprehensive mixed-use scheme has been granted planning permission comprising: a supermarket (7,756 sq m, 83,455 sq ft); retail units (6,532 sq m, 68,348 sq ft); restaurants and bar (7,724 sq m, 83,110 sq ft); health club and swimming pool (4,494 sq m, 48,355 sq ft); night club (3,718 sq m, 40,006 sq ft); casino (2,404 sq m, 25,867 sq ft); offices (12,620 sq m, 135,791 sq ft); and a car park (500 spaces).
- A Fund acquired part of the site (in a larger portfolio acquisition) at a book (accounting) cost of £8m, reflecting the development potential. It also inherited option agreements with other adjoining landowners in order to assemble the entirety of the site, which would result in a total site acquisition cost of £12.75m, thereby enabling the implementation of a comprehensive scheme.
- The costs of holding the site and keeping the options open with the other landowners are £150,000pa. Income from a car park on the site is £50,000pa. Therefore, net outgoings are £100,000pa (totaling £500k over 5 years, that is, this is assumed to be an intrinsic sunk cost in developing the site).
- The Council wishes to see the site comprehensively developed for the scheme and have granted permission. They also have a long-held objective of developing a sports and entertainment arena in the center of Croydon. Under an agreement with the investor in conjunction with granting the planning permission, the Council has said it would acquire the land at OMV (i.e., equivalent to the book cost) at any time up to 5 years from grant of planning permission should the investor wish to sell and not implement the scheme. Thereafter, the Council would acquire the site using CPO powers (a statutory procedure) if comprehensive development has not been started. The case for granting a CPO is believed to be given, among other reasons, due to the fragmented ownership and that this high-profile site has lain undeveloped for many years. Compensation from the Council to the Fund in acquiring the site via a CPO based on a *no scheme* world (i.e., ignoring any development potential) has been calculated at £5m.

Stage I: Mapping the Problem

Three basic real options were identified that conveyed *flexibility* in terms of optionality in real estate development. They were the option to abandon (i.e., sell), the option to defer investment, and the option to execute (i.e., implement the development). Any of these should be exercised prior to the expiration of 5 years given that the site would be compulsorily acquired at what the Fund estimated as being at subbook value under a CPO. In addition to these options, the option to alter the planning permission subject to market circumstances could also be added. While this would often occur in practice, it is not examined in this instance. The optionality of achieving an optimal tenant mix could also be considered.

As indicated earlier, these options are American (two calls and one put), although the decision just prior to the expiration of 5 years or the CPO could be considered a European put and therefore should be calculated as such.

The Croydon market was considered uncertain in terms of occupier requirements and rental levels, which were sensitive to general real estate market movements for both offices and retail. The ability to attract a supermarket operator and a major office pre-let were seen as key prerequisites prior to implementation of construction. The scheme would not be developed speculatively.

An ROA strategy matrix was prepared. Table 14.1 provides a simplified summary. It is evident from Table 14.1 that even in applying a qualitative

TABLE 14.1 ROA Development Strategy Matrix

Strategy/ Approach	Type of Development	Market Factors	Planning Issues	Timing	Embedded Option Appraisal
Pessimistic	Comprehensive Development	Poor office market; uncertain retail requirements	Reduce office content; reconfigure retail	3–5 yrs	Defer or sell
Cautious		Occupiers require 50% of offices; anchor retail tenant but at low rent gain	Consider phasing offices and retail (review planning obligations)	2–4 yrs	Defer or develop/ expansion option
Optimistic		Major office pre-let; quality anchor retailers secured; demand is high for all uses in the scheme	Consider increasing office content	1–3 yrs	Develop and expansion option

analysis, values may evolve asymmetrically. There could be a considerable upside relative to the downside. It was a characteristic of the Croydon office market, for example, that other competitor office schemes if implemented could encourage office sector activity and upward pricing of space with a high probability of occupier relocations. In this instance the investor did not have other real estate holdings in the town center. If the investor did, implementation of the scheme may also be considered a strategic (growth) option and could be analyzed as such.

Stage II: Base-Scoping Approach

A cash-flow residual development appraisal was produced, with key value drivers of the scheme being the supermarket and office components accounting for 47.15 percent of the expected capital value of the entire project. An overall blended yield of 7.8 percent was expected, which in market terms was considered cautious. An office rent of £215 per sq m (£20 per sq ft) was applied, although this was considered to have underperformed London's (and United Kingdom) office growth as illustrated in the two graphs in Figure 14.1. Total office returns also underperformed London (and the United Kingdom), which is in line with historic patterns for Croydon.

Costs comprised land acquisition, construction, professional fees, other agents' fees and costs, and finance (rolled up interest on costs). Land and construction costs excluding profit totaled £90.48m. The gross development value (GDV) of the scheme was £105.76m. It was considered by the investor that, for a project of this scale, a developer's profit on cost of 17.5 percent would be required (although profit on land was acceptable at 10 percent). The scheme on this basis outlined previously was marginally producing a total profit of £15.28m; in other words, a deterministic (NPV) measure of development profit. The next stage was to consider the project risks in a state without strategic flexibility.

A Monte Carlo simulation analysis was undertaken based on key input variables of supermarket and office rents and yields and office construction costs (a fuller analysis with other variables was initially undertaken and then narrowed down to key variables together with preliminary sensitivity and scenario analysis). The results are shown in the frequency chart in Figure 14.2.

Figure 14.2 shows a mean total profit return of £13.7m (90 percent certainty range of £8.3m to £19.0m) against a minimum required return of £14.7m (assuming 10 percent and 17.5 percent profit on land and construction cost, respectively). These returns can be compared with the ROA and explanation that incorporate a simulation of the option values in Figure 14.7 and Table 14.3, which appear in a later section. It should be noted that the project risk testing and use of simulation analysis, as illustrated earlier, is in itself a complex area, as highlighted earlier in this book.

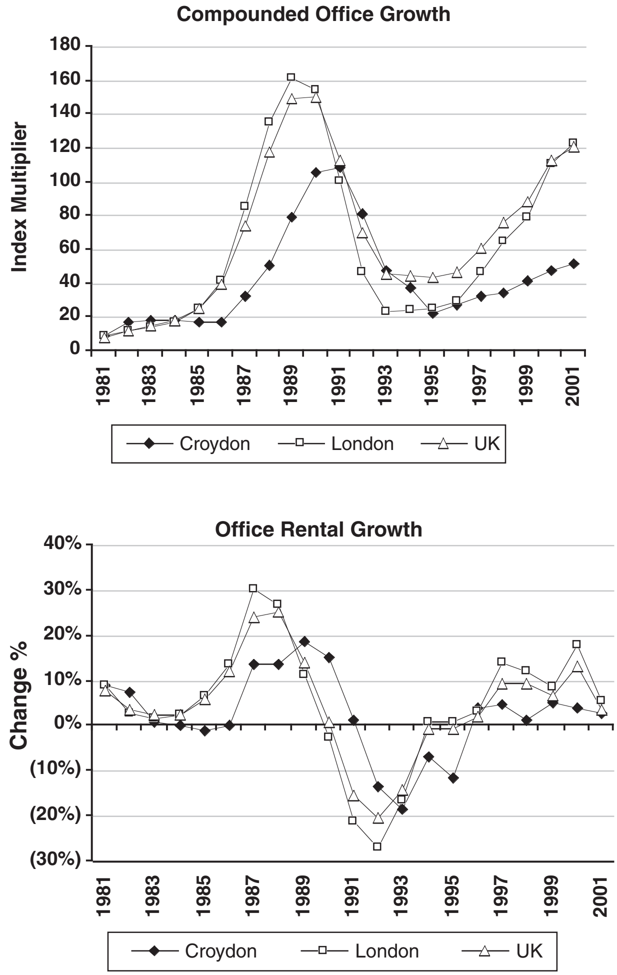


FIGURE 14.1 Croydon office rental and compounded growth. (*Source:* Data from IPD 2001)

Stage III: Internal and External Uncertainty Inputs

The base scoping provided a useful measure of the financial internal uncertainties and their interdependencies. In addition, it was necessary to regard specialist reports concerning construction constraints, cost variables, and programming. These also aided the simulation analysis in Stage II.

An ROA requires an assessment of volatility, a key input into the risk-neutral framework of real options pricing. In this instance, state pricing was

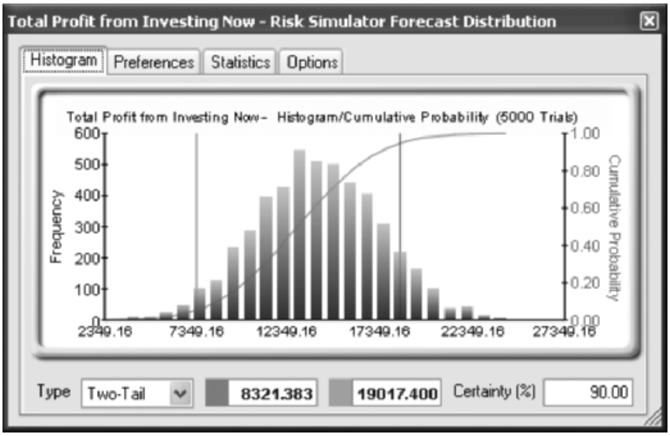


FIGURE 14.2 Base scoping Monte Carlo analysis.

also used. An assessment of the magnitude of the upside and downside within an underlying lattice in order to capture the likely asymmetry of the Croydon market was therefore undertaken.

As volatility is key to ROA, research and subsequent analysis are critical in obtaining suitable input data and then reviewing the resultant computations in Stage V. Indexes, as outlined later, are based on professional valuations as opposed to market transactions. Academic papers have highlighted the potential for what is known as valuation "smoothing" within the indexes with the result that volatility of real estate may be understated. Various techniques and data sources have been used for backing out true, historic, implied, and expected volatility in real estate over alternative time frames. However, this remains a significant area of research. The following approach has been simplified for practical reasons in obtaining appropriate volatility rates for this case study.

The U.K. Investment Property Databank (IPD) data on office and retail rental growth and total returns for Croydon, London, and the United Kingdom between 1981 and 2002 were analyzed. As investment performance is judged on total returns, these volatility figures were used with respect to the underlying asset value. Volatility of total returns for office and retail for three periods—1981–2002(1); 1991–2002(2); and 1995–2001(3)—are shown in Figure 14.3. Both graphs show volatility decreasing over the three periods from a range of 8.6 percent to 12.1 percent (offices) and 6.4 percent to 8.7 percent (retail) to 2.4 percent to 3.3 percent (offices) and 1.15 percent to 3.4 percent (retail). These appear to be low volatility rates compared to empirical research.

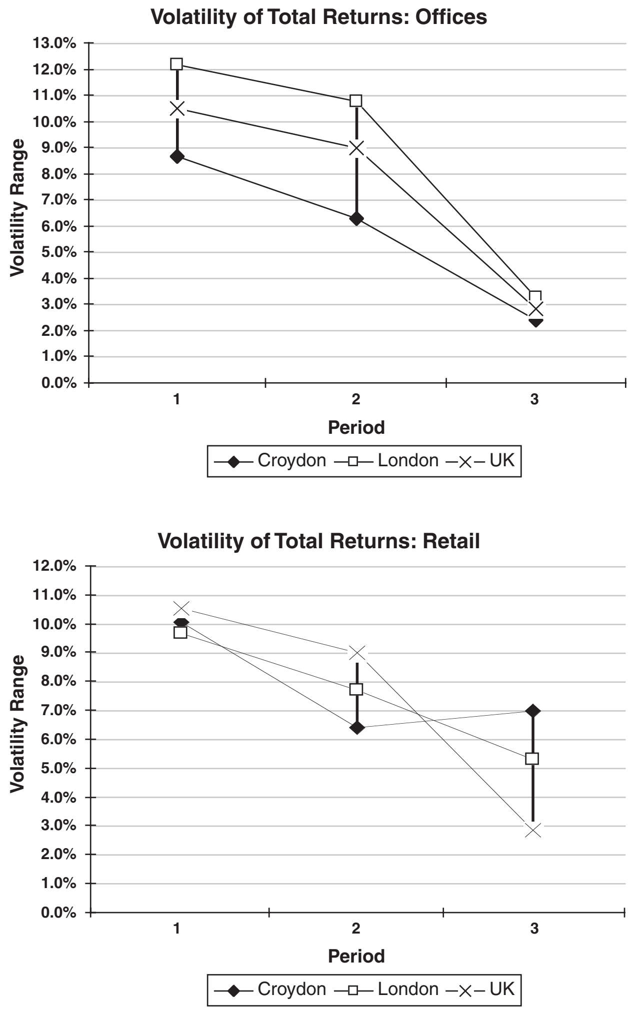


FIGURE 14.3 Croydon retail and office volatility of total returns. (*Source:* Data from IPD 2001)

Another way of considering the volatility over this period for offices and retail is on a 5-year rolling basis as shown in the two charts in Figure 14.4.

From Figure 14.4 we see that the Croydon office market showed an average volatility of 8.95 percent (range 2.2 percent to 14.7 percent), which was below both London (average 11.39 percent, range 4.1 percent to 24.1 percent) and the United Kingdom (average 10.12 percent, range 2.6 percent to 10.9 percent). For retail (except in Croydon) the volatility levels were

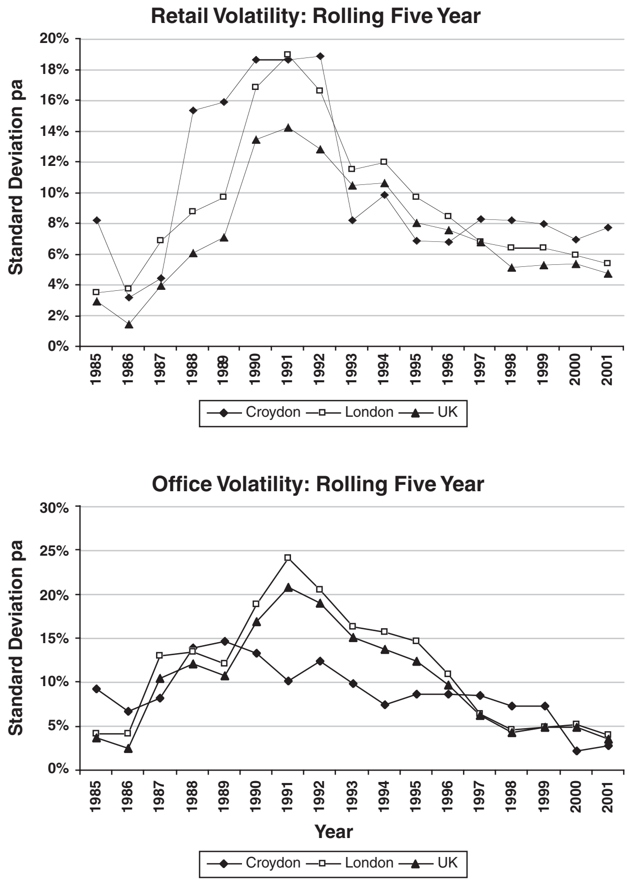


FIGURE 14.4 Croydon retail and office returns—5-year rolling volatility. (*Source:* Data from IPD 2001)

generally lower than for offices, with the Croydon market showing an average of 10.27 percent (range 3.2 percent to 18.9 percent) compared with a London average of 9.29 percent (range 3.5 percent to 19 percent) and the United Kingdom average of 7.46 percent (range 1.5 percent to 14.3 percent).

It is necessary for the underlying asset to arrive at a single volatility, that is, combining retail and offices. Further research and analysis in practice was undertaken, including cross correlations. For the purposes here, a figure of 10 percent with an analysis range of between 5 percent and 35 percent is utilized, taking account of sector empirical studies and desmoothing of base indexes.

So far as the price probability falls under the ROA analytical approach state pricing, this has regard to compounded growth in capturing the asymmetry of future underlying asset changes. Again, further research in practice was undertaken. Indeed, an alternative approach in option pricing would be via a jump-diffusion whereby an initial jump (i.e., upside) could be followed by a reversion to appropriate volatility levels. Nonrecombining lattices or multiple recombining lattices with changing volatilities could also achieve similar results. For state pricing, the upstate was assumed at 15 percent and downstate 5 percent. See Johnathan Mun's *Real Options Analysis, Second Edition* (Wiley, 2005) for technical details.

So far as costs were concerned, cost inflation was set at 5 percent and cost volatility at 5 percent. The latter was considered low in comparison to empirical examples and therefore was analyzed within a range of 5 percent to 25 percent. U.K.-published construction cost indexes have been criticized as not reflecting the true volatility found in the sector. This criticism has again led to other alternative measures and proxies being sought and analyzed, including traded call options of construction companies.

Stage IV: Real Options (Quantitative) Analysis

The three lattice approaches together with the inputs and assumptions outlined earlier were computed. The cost of implementation input excluded profit on cost and land in order to directly compare the option price to development profit. The value input was that derived from the deterministic appraisal. Under each approach, the lattices were as follows:

- An underlying asset pricing lattice, the price evolution.
- An underlying cost lattice, the cost growth or evolution.
- The value of exercising the development, in simple terms the NPV in each moment of time of making an investment.
- A valuation lattice, where the value would be the maximum of price less cost; the option to defer less the intrinsic sunk costs; or the offer to be acquired by the Council. The termination boundary (year 5) would be the maximum of the underlying price less costs or the offer to be acquired by the Council.
- A decision lattice, which was based on the valuation lattice in determining at each node whether to defer, sell, or develop.

Option values were calculated under each of the three approaches, which were then compared to the development profit of the deterministic approach, as shown in Table 14.2. In each case the value (profit) of the option to defer (i.e., now or later) is higher than the current or expected profit of investing immediately. The difference in the real option values results from the evolution of the lattice and risk-neutral pricing of each approach.

TABLE 14.2 A Comparison of Real Option Values with NPV

NPV (£m)	ROA Binomial (£m)	ROA State Pricing (£m)	ROA Binomial (Dual Volatility) (£m)
15.28[a]	18.13	18.09	23.77
Additional Value Created by ROA	2.85	2.81	8.49

[a]This amount represents the total profit of investing now of which £14.7m would be the minimum required return.

Stage V: Explanation

The option price takes into account all possible future outcomes under the three ROA approaches that were not captured by the deterministic analysis. It was, however, necessary to consider the sensitivity of the inputs, particularly with respect to volatility (price and cost) and price probabilities under state pricing as well as the impact on the decision lattice at the different nodes. The decision lattices in Figure 14.5 (with time in years in bold) are set out for comparative purposes.

Taking an overview with regard to all of the approaches, development should probably be deferred in years 1 and 2; deferral or selling were the dominant options in year 3; and development should only probably be envisaged in years 4 or 5. This scheme essentially provided an analytical underpinning for a professional judgment and decision framework. The surface graphs in Figure 14.6 illustrated the sensitivity for each approach. Figure 14.6 clearly indicated the effect and interaction of volatility on the option price (OP), which again emphasized the importance attached to establishing base volatility inputs as discussed earlier in Stage II. This analysis in practice was analyzed and reported on further. A Monte Carlo analysis of each option price was undertaken and the frequency charts are set out in Figure 14.7 together with a certainty level of 90 percent. These charts can be compared to the base-scoping frequency chart (Figure 14.2) and illustrate the narrowing (particularly with state pricing) of the risk structure and higher average return.

It was notable that the risk structure range's downside of the three approaches was relatively similar, being between £16.2m and £18.6m (see Table 14.3). In this particular instance, the downsides provided useful benchmarks to the minimum required return of £14.7m under an NPV approach, as an alternative measure to comparing average returns. Notwithstanding this NPV result, the upsides under the three approaches were significant.

The investor, as a result of an ROA, could clearly form a strategy in terms of optimal timing or whether to invest at all. The flexibility of this

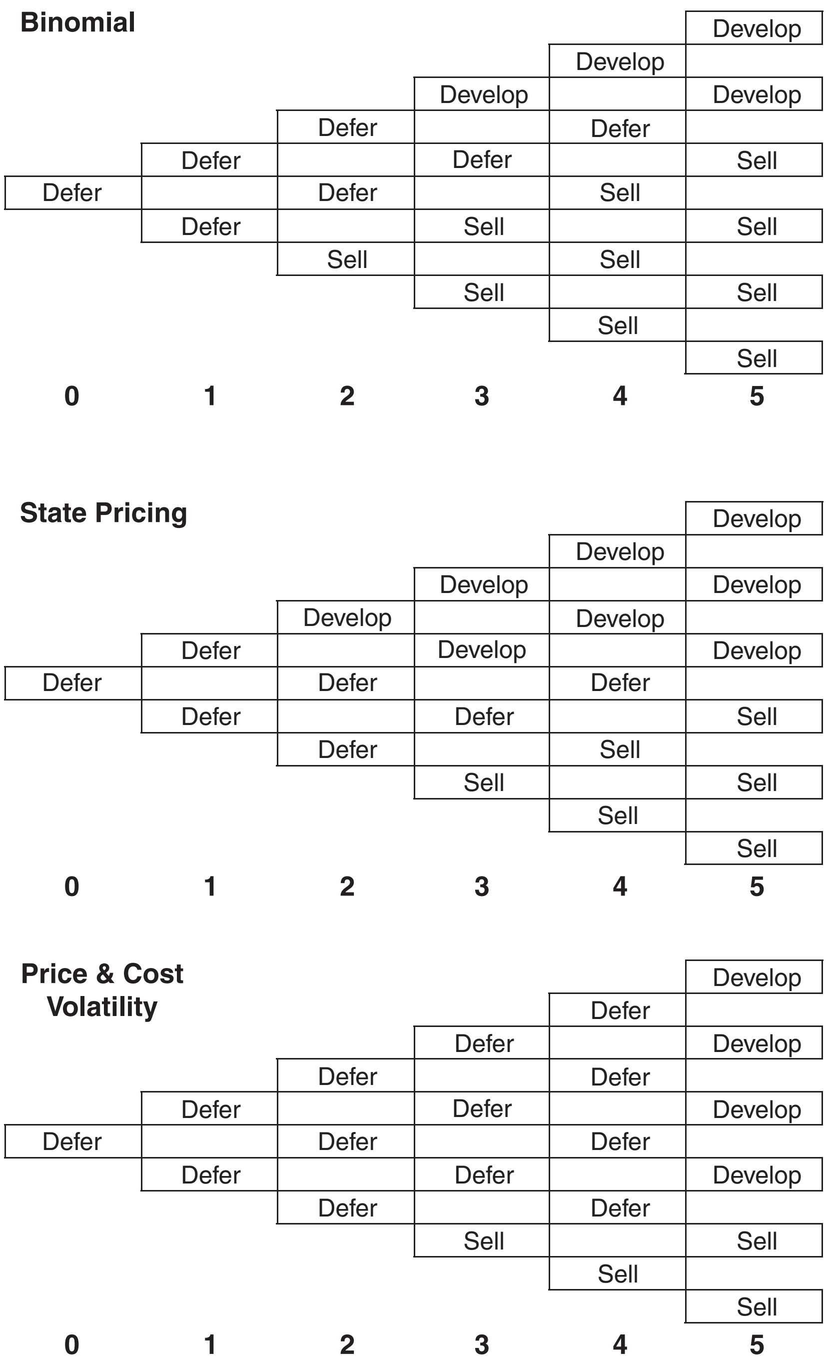


FIGURE 14.5 Binomial lattices.

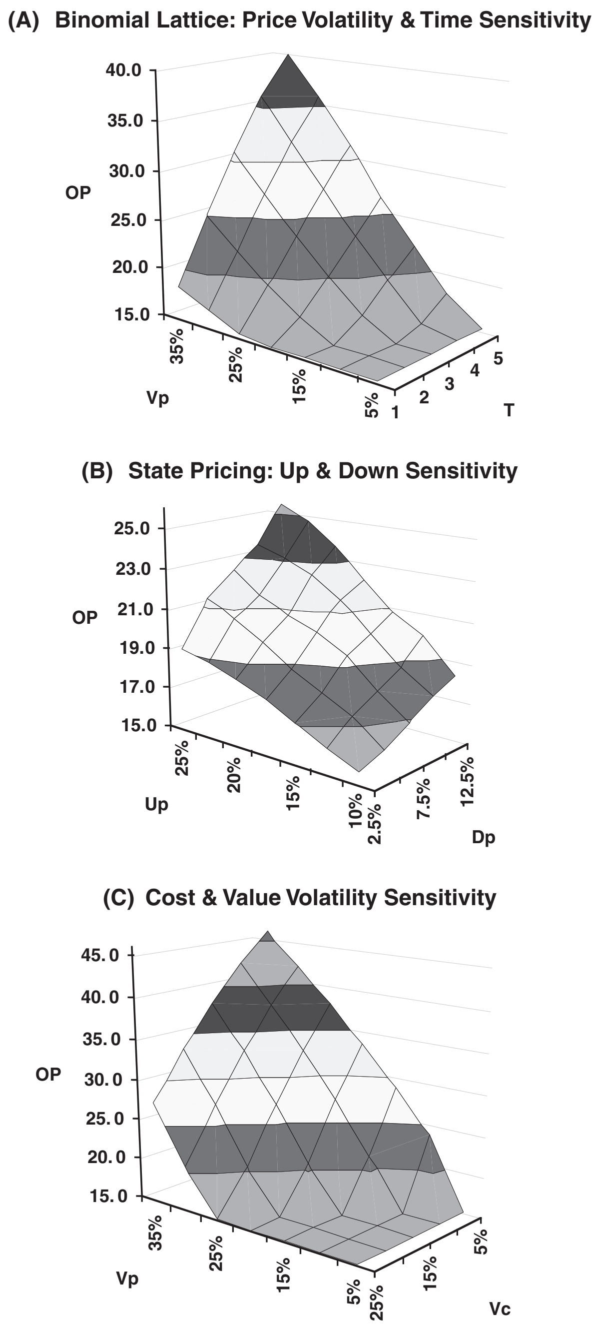


FIGURE 14.6 Croydon ROA sensitivity graphs.

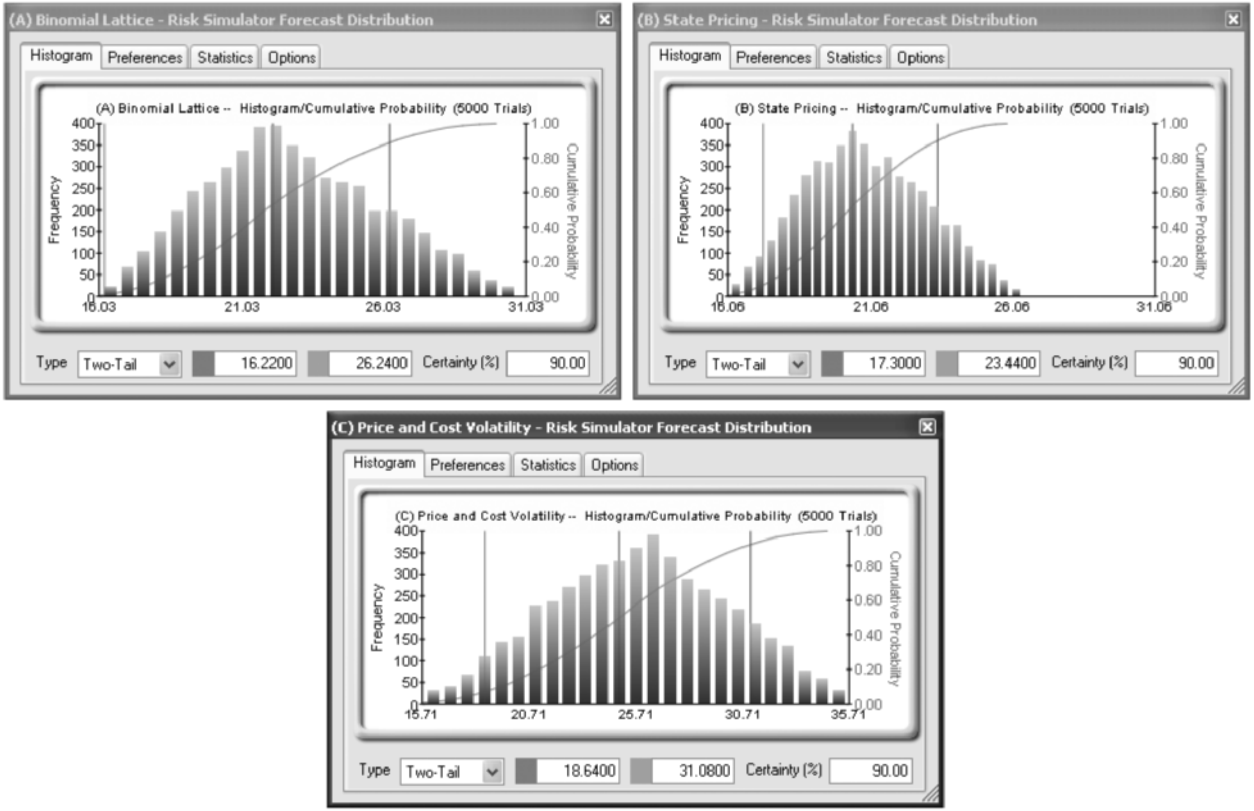


FIGURE 14.7 ROA Monte Carlo application.

TABLE 14.3 Simulated NPV and Option Values Croydon (Average and Range)

	Average Return (£m)	Risk Structure Range 90% (£m)	Percentage Above Required Return (£m)
NPV	13.7	8.3–19.0	(6.8)
Binomial Lattice	21.1	16.2–26.2	43.5
State Pricing	20.6	17.3–23.4	40.0
Binomial (Cost/Price Volatility)	25.1	18.6–31.1	70.7

decision created additional value over and above a conventional valuation of the development. This additional value would perhaps be incorporated within a price, if the investor were to dispose of the opportunity to a third party at the beginning of the period.

The real option paradigm when applied to real estate potentially highlights, on one hand, the seemingly intuitive action of investors and, on the other hand, undervalued investment opportunities and suboptimal decisions. As such the ROA, as illustrated previously, therefore provides another approach and valuable layer to the risk analysis and potential returns of real estate investment and development.

Notes

The following papers provide further reading on the subjects of investment risk, volatility measures, and real options in real estate development.

Brown, G., and G. Matysiak. *Real Estate Investment, A Capital Market Approach*. London: Financial Times Prentice Hall, 2000.

Grenadier, S. "The Strategic Exercise of Options: Development Cascades and Overbuilding in Real Estate Markets." *Journal of Finance* 51, no. 5 (1996): 1653–1679.

Quigg, L. "Empirical Testing of Real Option-Pricing Models." *Journal of Finance* 68, no. 2 (1993): 621–639.

Sing T. "Optimal Timing of Real Estate Development under Uncertainty." *Journal of Property Investment & Finance,* Special Issue: Real Options, 19, no. 1 (2001): 35–52.

Titman, S. "Urban Land Prices under Uncertainty." *The American Economic Review* 75, no. 3 (1985): 505–514.

Ward C. "Arbitrage and Investment in Commercial Property." *Journal of Business & Accounting* 9, no. 1 (1982): 93–108.

Williams, J. "Real Estate Development as an Option." *Journal of Real Estate Finance and Economics* 4, no. 2 (1991): 191–208.

CASE STUDY: USING STOCHASTIC OPTIMIZATION AND VALUATION MODELS TO EVALUATE THE CREDIT RISK OF CORPORATE RESTRUCTURING

This business case is contributed by Professor Morton Glantz. Professor Glantz is on the finance faculty of Fordham Graduate Business School in New York. He is widely published in financial journals and has authored a number of books, including Optimal Trading Strategies, Managing Bank Risk, Scientific Financial Management, *and* Loan Management Risk.

Companies restructure their product mix to boost sales and profits, increase shareholder value, or to survive when the corporate structure becomes impaired. In successful restructurings, management not only actualizes lucrative new projects, but abandons existing projects when they no longer yield sufficient returns, thereby channeling resources to more value-creating uses.

At one level, restructuring can be viewed as changes in financing structures and management. At another level, restructuring may be operational—in response to production overhauls, market trends, technology, and industry or macroeconomic disturbances. It is often the essence of strategy formulation—that is, management's response to changes in the environment to creatively deploy internal resources—that improves the firm's competitive position. Indeed, changing operating and financial structures in pursuit of a long-run strategy is a key corporate goal—the most direct path to shareholder value.

For banks called on to finance corporate restructurings, things are a bit different. For example, most loans provide a fixed return over fixed periods that are dependent on interest rates and the borrower's ability to pay. A good loan will be repaid on time and in full. It is hoped that the bank's cost of funds will be low, with the deal providing attractive risk-adjusted returns. If the borrower's business excels, *the bank will not participate in upside corporate values* (except for a vicarious pleasure in the firm's success). However, if a borrower ends up financially distressed, lenders share much, perhaps most, of the pain.

Two disparate goals—controlling default (credit) risk, the bank's objective, and value maximization, a traditional corporate aspiration—are often at odds, particularly if borrowers want term money to finance excessively aggressive projects. In the vast majority of cases of traditional credit analysis, where the spotlight focuses on deterministically drawn projections, hidden risks are often exceedingly difficult to uncover. Devoid of viable projections, bankers will time and again fail to bridge gaps between their agendas and client aspirations.

This case study offers ways for bankers to advance both their analytics and communication skills—senior bank officials and clients alike to "get the deal done" and ensure risk/reward agendas are set in equilibrium. Undeniably, the direct way to achieve results is to take a stochastic view of strategic plans rather than relying inappropriately on deterministic base case or conservative scenarios. Let us start with the following fundamentals:

- Stochastically driven optimization models allow bankers to more realistically represent the flow of random variables.
- In negotiating restructuring loans, borrowers (and bankers) can determine under stochastic assumptions optimal amounts to invest in or borrow to finance projects.
- McKinsey & Company, Inc.,[1] suggests that business units should be defined and separated into lines of business. Business units should be broken down into the smallest components and analyzed at the base level first.
- Consolidating financials, rather than consolidated reports, should be used to perform business-unit valuations.
- Knowing the market value and volatility of the borrower's assets is crucial in determining the probability of default.
- A firm's leverage has the effect of magnifying its underlying asset volatility. As a result, industries with low-asset volatility can take on larger amounts of leverage, whereas industries with high-asset volatility tend to take on less.
- After restructuring is optimized at the unit stage, unit level valuations are linked to the borrower's consolidated worksheet to process corporate valuations.

The Business Case

Consider the data in Excel spreadsheets depicted in Figures 14.8, 14.9, and 14.10. The worksheets depict management's original restructuring plan. ABC Bank is asked to approve a $3,410,000 loan facility for the hypothetical firm RI Furniture Manufacturing LTD. Management wants to restructure four of its operating subsidiaries. In support of the facility, the firm supplied the bank with deterministic base case and conservative consolidating and consolidated projections—income statement, balance sheet, and cash flows.

The deterministic or static forecasts tendered the bank limited the variability of outcomes. From a banker's perspective it is often difficult to single out which of a series of *strategic options* the borrower should pursue if the bank fails to understand differences in the range and distribution shape of

	Distribution	Operating Profit Margin Range	Operating Profit Margin Most Likely
All Weather Resin Wicker Sets	Triangular	5.5% – 12.6%	11.0%
Commuter Mobile Office Furniture	Triangular	6.5% – 8.7%	7.5%
Specialty Furniture	Triangular	0.5% – 5.3%	4.7%
Custom Built Furniture	Uniform	3.3% – 6.6%	None

FIGURE 14.8 Distributional assumptions.

possible outcomes and the most likely result associated with each option. Indeed an overly aggressive restructuring program might reduce the firm's credit grade and increase default probabilities. We will not let this happen. Undeniably, this deal deserves stochastic analytics rather than a breadbasket consisting of passé deterministic tools.

From deterministic consolidating projections, bankers developed a stochastic spreadsheet depicted in Figure 14.10. This spreadsheet included maximum/minimum investment ranges supporting restructuring in each of four product lines. Using optimization along with the deterministic McKinsey DCF Valuation 2000 Model, the firm's bankers came up with a stochastic solution. On a unit level, they developed a probability distribution assigned to each uncertain element in the forecast, established an optimal funding array for the various business combinations, and held cash-flow volatility to acceptable levels, preserving the credit grade (again at the unit level). Finally, the last optimization (worksheet) was linked to the consolidating/consolidated DCF valuation worksheet(s). The firm's bankers then determined postrestructuring equity values, specific confidence levels, and probabilities that asset values fall below debt values.

Product Line	Lower Bound	Upper Bound
All Weather Resin Wicker Sets	1,000,000	1,250,000
Commuter Mobile Office Furniture	600,000	1,000,000
Specialty Furniture	570,000	1,100,000
Custom Built Furniture	400,000	900,000

FIGURE 14.9 Investment boundaries.

	B	C	D	E	F
1	**RI Furniture Co. Limited: Strategic Plan**				
2					
3		*Annual*	*Lower*	*Upper*	
4	*Proposed New Product Lines*	*operating return*	*bound*	*bound*	
5	All Weather Resin Wicker Sets	9.7%	$1,000,000	$1,250,000	
6	Commuter Mobile Office Furniture	7.6%	$600,000	$1,000,000	
7	Specialty Furniture	3.5%	$570,000	$1,100,000	
8	Custom Built Furniture	5.0%	$400,000	$900,000	
9					
10					
11		*Amount*			
12	*Decision variables*	*invested*	Constraint		
13	All Weather Resin Wicker Sets	$1,125,000	Decision Variables		
14	Commuter Mobile Office Furniture	$800,000	prior to optimization		
15	Specialty Furniture	$835,000			*Total amount*
16	Custom Built Furniture	$650,000			*invested*
17	*Total expected return*	$231,058	Objective		$3,410,000
18	*(Annual operating return X Amount invested)*				

FIGURE 14.10 Borrower's original strategic restructuring plan (reworked by the bank in a stochastic mode, not yet optimized).

Business History

RI Furniture started operations in 1986. The firm manufactures a full line of indoor and outdoor furniture. Operating subsidiaries targeted for restructuring, depicted later, represent approximately 65 percent of consolidated operations.

- *All Weather Resin Wicker Sets*. This furniture comes with a complete aluminum frame with handwoven polypropylene resin produced to resist weather. *Operating profit margin distributions and investment ranges for each subsidiary are shown in Figures 14.8 through 14.10.*
- *Commuter Mobile Office Furniture*. The commuter rolls from its storage location to any work area and sets up in minutes. It integrates computer peripherals (monitor, CPU tower, keyboard, and printer) in a compact, secure mobile unit.
- *Specialty Furniture*. After restructuring, this business segment will include production of hotel reception furniture, cafe furniture, canteen furniture, restaurant seating, and banqueting furniture.
- *Custom-Built Furniture*. Furniture will be custom built in the firm's own workshop or sourced from a host of reputable manufacturers both at home and abroad.

The analysis was run by placing a constraint on $3,410,000 investment—that is, the bank's facility cannot exceed $3,410,000. Later we place an additional constraint: the forecast variable's volatility. From the information in Figures 14.8 and 14.9, the bank developed the spreadsheet depicted in Figure 14.10.

Using optimization, a constraint on investment/loan facility was entered:

All Weather Resin Wicker Sets + *Commuter Mobile Office Furniture* + *Specialty Furniture* + *Custom Built Furniture* <=3410000.

Note that investment falls to within the constraint boundary, while expected return increased.

Simulation statistics reveal that volatility of the expected return (the forecast variable), as measured by the standard deviation, was $20,000. Again, *volatility of operating results affects the volatility of assets*. This point is important. Suppose we determine the market value of a corporation's assets as well as the volatility of that value. Moody's KMV demonstrates that volatility measures the propensity of asset values to change within a given time period. This information determines the probability of default, given the corporation's obligations. For instance, KMV suggests that if the current asset market value is $150 million and a corporation's debt is $75 million and is due in 1 year, then default will occur if the asset value turns out to be less than $75 million in 1 year. Thus, as a prudent next step, bankers discuss the first optimization run (Figure 14.11) with management on three levels: (1) maximum expected return, (2) optimal investments/loan facility, and

	A	B	C	D	E	F
10						
11			*Amount*	Constraint		
12		*Decision variables*	*invested*			
13		All Weather Resin Wicker Sets	$1,247,100			
14		Commuter Mobile Office Furniture	$993,671	*Decision Variables prior to optimization*		
15		Specialty Furniture	$570,000			*Total amount*
16		Custom Built Furniture	$598,998			*invested*
17		*Total expected return*	$245,757	Objective		$3,409,769
18		*(Annual operating return X Amount invested)*				
19						

Total Expected Return - Risk Simulator Forecast Distribution

Histogram | Preferences | Statistics | Options

Statistics	Result
Number of Trials	1000
Mean	248378.1019
Median	248845.3630
Standard Deviation	20278.0716
Variance	411200188.4180
Average Deviation	15951.5389
Maximum	314095.8767
Minimum	176695.9212
Range	137399.9555
Skewness	-0.0396
Kurtosis	0.2564
25% Percentile	234474.0950
75% Percentile	261455.6607
Error Precision at 95% confidence	0.5060%

FIGURE 14.11 Run Two optimization results.

(3) volatility of expected return. If volatility is unacceptable, the standard deviation must be reduced to preserve credit grade integrity. We assume the bank requires project standard deviation to be equal to or below $17,800.

The final simulation shown in Figure 14.12 produced an optimization that reconciled both risk/reward agendas discussed earlier. The loan facility effectively reduces to (optimized) $3,331,102, and because the firm requires less money, financial leverage improves. We note that $227,889 is the maximized expected return, lower than the $245,757 produced with no volatility constraint—lower risk reduces rewards.

The story does not end here; our analysis up to now was restricted to the unit level—that is, business segments involved in the restructuring. While the spreadsheet in Figure 14.12 worked its stochastic wonders, it *must now link to consolidating and consolidated discounted cash-flow (DCF) valuation worksheets*. Consolidated DCF valuations provide a *going-concern* value—the value driven by a company's future economic strength. RI Furniture value is determined by the present value of future cash flows for a specific forecast horizon (projection period) plus the present value of cash flow *beyond* the forecast horizon (residual or terminal value). In other words, the firm's value depends on cash-flow potential and the risks (threats) of those future cash flows. These perceived risks or threats help define the discounting factor used to measure cash flows in present value terms. Cash flow

	B	C	D	E	F
10					
11		*Amount*	Constraint		
12	*Decision variables*	*invested*			
13	All Weather Resin Wicker Sets	$1,000,000			
14	Commuter Mobile Office Furniture	$993,225	Decision Variables prior to optimization		
15	Specialty Furniture	$723,457			*Total amount*
16	Custom Built Furniture	$614,420			*invested*
17	*Total expected return*	$227,889	Objective		$3,331,102
18	*(Annual operating return X Amount invested)*				
19		**Expected**	**Total**		
20	**Summary**	**Return**	**Investment**	**Standard Deviation**	
21	**Borrower's Original Projections**	$231,058	$3,410,000	n/a	
22	**Run One: Original Projections Optimized**	$245,757	$3,409,769	$20,373	
23	**Run Two: Project Volatility Constraint**	$227,889	$3,331,102	$17,800	
24	**Run Two: Project Volatility Actual**			$17,701	
25	**Expected Return and Loan Reduction**	$17,868	$78,667		
26	***(Bank Requirement: Reduce Project Risk)***				
27					
28			**Run One**	**Run Two**	
29	**Investment (Loan Amounts)**	**Original Strategy**	**Optimized; No**	**Optimized; Risk**	
30		**Not Optimized**	**Risk Constraint**	**Constraint**	
31	**All Weather Resin Wicker Sets**	$1,125,000	$1,247,100	$1,000,000	
32	**Commuter Mobile Office Furniture**	$800,000	$993,671	$993,225	
33	**Specialty Furniture**	$835,000	$570,000	$723,457	
34	**Custom Built Furniture**	$650,000	$598,998	$614,420	
35	**Total**	$3,410,000	$3,409,769	$3,331,102	

FIGURE 14.12 Final optimization results.

depends on the industry and the economic outlook for RI Furniture's products, current and future competition, sustainable competitive advantage, projected changes in demand, and this borrower's capacity to grow in light of its past financial and operating performance. Risk factors that the firm's bankers will examine carefully include their borrower's financial condition; quality, magnitude, and volatility of cash flows; financial and operating leverage; and management's capacity to sustain operations on a profitable basis. *These primary attributes cannot be ignored when bankers determine distributions associated with assumption variables.*

Simulation and optimization embedded into powerful valuation models provides an intuitive advantage; it is a decidedly efficient and precise way to get deals analyzed, done, and sold.

CASE STUDY: REAL OPTIONS AND KVA IN MILITARY STRATEGY AT THE UNITED STATES NAVY

This case was written by Lieutenant Commander Cesar Rios in collaboration with Dr. Tom Housel and Dr. Johnathan Mun. Lieutenant Commander Rios is an intelligence officer for the U.S. Navy assigned to the Third Expeditionary Strike Group in San Diego, California. Dr. Tom Housel is a professor of Information Sciences at the Naval Postgraduate School in Monterey, California. Please contact Dr. Housel with any questions about the case at tjhousel@nps.edu.

Millions of dollars are spent by the United States military for information technology (IT) investments on Quick Reaction Capability Information Warfare (IW) and intelligence collection systems. To evaluate and select projects yielding maximum benefits to the government, valuation tools are critical to properly define, capture, and measure the total value of those investments. This case study applies Knowledge Value-Added (KVA) and Real Options valuation techniques to the Naval Cryptologic Carry-On Program (CCOP) systems used in the intelligence collection process, with particular focus on human capital and IT processes. The objective is to develop a model and methodology to assist in the budgeting process for IW systems. The methodology had to be capable of producing measurable objectives so existing and future CCOP systems could be evaluated.

The Challenge

The Chief of Naval Operations directed its CCOP Office to focus on three goals for fiscal year 2005: efficiencies, metrics, and return on investment.

Given this mandate, CCOP Program Manager Lieutenant Commander (LCDR) Brian Prevo had the difficult choice of how much funding to allocate among the 12 IW CCOP systems currently in his portfolio. Should he merely allocate an equal amount of continuous funding? Should he ask which ones needed the most funding to continue or upgrade? Should he ask the users which ones they preferred? To make appropriate budget decisions, LCDR Prevo had to analyze the operating performance of each CCOP program by developing metrics, measuring efficiencies, and calculating the return on investment. Moreover, he had to identify which investment options supported the United States Navy's Global Intelligence, Surveillance, and Reconnaissance (ISR) mission. LCDR Prevo teamed with researchers at the Naval Postgraduate School (NPS). He enlisted Professor Thomas Housel and Professor Johnathan Mun at NPS's Graduate School of Operations and Information Sciences to identify valuation techniques to help manage his CCOP portfolio. Prevo also sought the aid of NPS student LCDR Cesar Rios, a Naval Cryptologist and Information Warfare Officer. Rios had operated CCOP systems and other IW systems while conducting ISR missions from various Navy platforms, including ships and aircraft. As the team leader and subject matter expert, LCDR Rios worked with Dr. Housel and Dr. Mun to conduct the analysis required to make the optimal portfolio management decision in his CCOP strategies.

Background

Intelligence is a critical component of U.S. security strategy. It is the first line of defense against threats poised by hostile states and terrorists, according to the National Security Strategy (NSS) of the United States.[2] After the tragic events of September 11, a new world emerged where intelligence techniques from the Cold War era were inadequate to meet the new and complex security threats to the United States. Several initiatives were launched to transform the country's intelligence capabilities to keep pace with emerging threats, including:

- Establishing a new framework for intelligence warning providing seamless and integrated warning across the spectrum of threats facing the nation and its allies.
- Developing new methods for collecting information to sustain intelligence advantage.
- Investing in future capabilities while working to protect them through a more vigorous effort to prevent the compromise of intelligence capabilities.
- Collecting intelligence against the terrorist danger across the government with all-source analysis.[3]

Expenditures on U.S. intelligence activities are estimated at $40 billion annually and a significant amount of that total is spent on ISR activities. The ISR are the systems that gather, process, and disseminate intelligence. The ISR systems cover a multitude of systems and programs for acquiring and processing information needed by national security decision makers and military commanders. The ISR systems range in size from small, hand-held cameras to billion dollar satellites. Some ISR programs collect basic information for a wide range of analytical products, whereas others are designed to acquire data for specific weapons systems. Some are "national" systems, collecting information for government agencies, whereas others are "tactical" systems intended to support military commanders on the battlefield. The ISR programs are currently grouped into three major categories: the National Intelligence Program (NIP), the Joint Military Intelligence Program (JMIP), and Tactical Intelligence and Related Activities (TIARA).

Most intelligence used by the military comes from the Defense Intelligence Agency (DIA), which produces some HUMINT, MASINT, and a large portion of the Department of Defense's (DoD) strategic, long-term analysis; the National Security Agency (NSA), which produces most SIGINT; and the National Imagery and Mapping Agency (NIMA), which produces most IMINT.[4] To a lesser extent, the military intelligence community also consists of the Central Intelligence Agency (CIA), State Department, Department of Energy, Department of Justice, and Department of Treasury.

Navy ISR

The Naval Transformation Roadmap of 2003 calls for the reengineering of maritime ISR to align with the DoD's 5000 Series and joint warfighting concepts. Goals are to redefine standards and metrics and ensure interoperability while providing the warfighter-required capabilities in a timely, cost-effective, and efficient manner. Maritime ISR lies at the core of the Naval Operational Doctrine and is an essential element in improving the speed and effectiveness of naval and joint operations. With today's security threats, it is necessary to expand the range of ISR options available to the commander and ensure decision superiority across the range of military operations in accordance with the NSS.

The Intelligence Collection Process (ICP) is the way tactical Navy ISR units of ships, aircraft, and other platforms complete intelligence requests. Once requests are received, human disciplines and IT technologies are used together to search, acquire, process, and report results back to tactical users (i.e., fleet staffs and strike groups), and national-level consumers (i.e., NSA). The generalized process is shown in Figure 14.13.

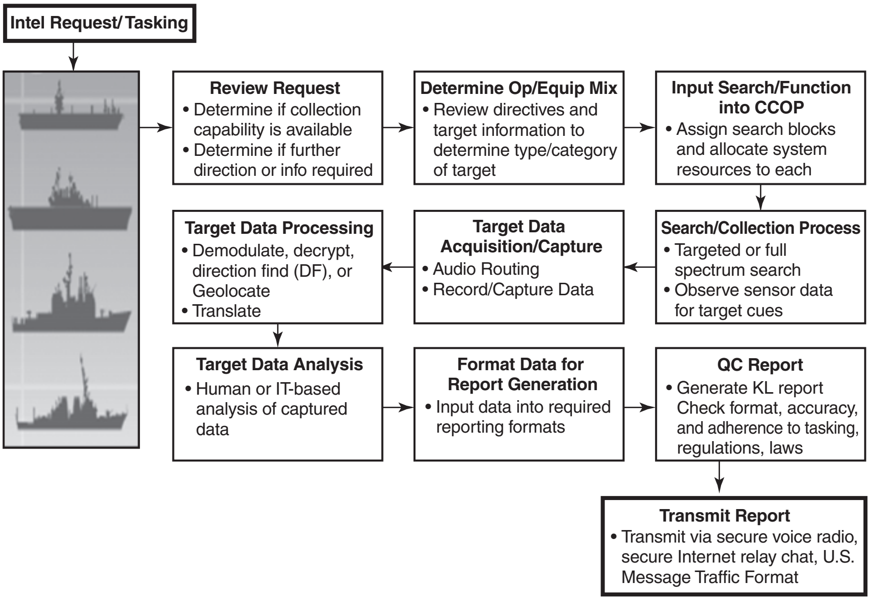


FIGURE 14.13 The intelligence collection process.

Each subprocess is further broken down into individual actions that may be required to perform the subprocess in the ICP. For example, the subprocess Target Data Processing can be broken down into a number of tasks:

1. Human-based (no automation required)
   a. Manual copy directly into report
   b. Human translation and processing
2. IT-based
   a. Direct transfer into report
      (1) Demodulate
         (a) All IT-based
         (b) Human enabled
      (2) Decrypt
         (a) All IT-based
         (b) Human enabled
   b. Direction finding
      (1) Automatic—Local line of bearing (LOB)
      (2) Human-enabled—Local LOB
      (3) Human-enabled—B-rep request
   c. Geolocation
      (1) Special processing

Established in 1994, CCOP developed state-of-the-art ISR capabilities for Combatant Command requirements for a quick-reaction surface, subsurface, and airborne cryptologic carry-on capability. Approximately 100 cryptologic-capable surface ships are currently in the U.S. Navy inventory. Each one is a potential user of carry-on equipment, along with numerous subsurface and air platforms. Although CCOP systems have broad scope and functions, basic capabilities include:

- Tactical surveillance, targeting.
- Passive detection, classification, tracking, enemy intent at extended range.
- Analysis tools allowing interpretation and reporting of the potential or known meaning of intercepted data.
- Correlation and tracking.

As part of the Advanced Cryptologic Systems Engineering program, CCOP utilizes commercial off-the-shelf (COTS) technology, government off-the-self (GOTS) technology, and modular, open systems architectures. COTS and GOTS technologies, when applied to ISR system functionalities, typically require various levels of integration to leverage on-board capabilities to provide system and mission management, product reporting, and data analysis support. COTS and GOTS also require some level of adaptation or modification to meet fleet requirements. Before deployment for operational use, systems must be systematically tested to ensure suitable and reliable operation. They must also be tested for network vulnerabilities (if connected to Navy or joint networks), and tested against joint interoperability requirements.

Valuation Techniques

Assessing information technology investments is a daunting challenge. Although several valuation methods are used to measure and justify IT investments, return on investment (ROI) is the most widely used metric to measure past, present, and potential future performance. Other techniques are used to measure the impact of IT on organizations at the corporate and subcorporate levels. Although approaches differ, the objectives are similar and that is to provide managers with metrics to measure tangible IT investments and intangible knowledge assets. Corporate-level approaches determine the contribution of both IT and knowledge assets on the overall performance of the organization. Subcorporate-level approaches look internally at the subprocesses involved in the production of organizational output and attempt to establish a measure for the benefits of knowledge and IT assets within each subprocess.

ROI in the Public Sector

ROI yields insights for managers and investors making high-level strategic business decisions, yet what if an organization does not produce measurable revenues such as the U.S. DoD? Traditional ROI metrics cannot measure the total value of IT investments made by public sector entities. When conducting an ROI analysis for the public sector, there are several considerations:

- Lack of measurable revenues and profits makes it challenging to determine the overall benefit stream produced by the organization.
- Concrete data is often difficult to collect amid an abundance of seemingly intangible soft data.
- ROI depends on costs and benefits; recipients of benefits or stakeholders are not easily identifiable because potential beneficiaries are program participants, managers of participants, program sponsors, or taxpayers.
- Certain government services are essential for the public good and must be provided, regardless of the accountability or cost.

Budgets of public sector organizations are under increased scrutiny, with stakeholders, managers, and taxpayers demanding higher levels of accountability and transparency of public investment. Compounding the problem further are increased regulations such as the Government Performance Results Act of 1993 (GPRA), requiring the establishment of strategic planning and performance measurement in programs for the accountability of their expenditures. These challenges have forced public sector organizations to adopt quantifiable methods to produce the required metrics for measuring the *total value* of services and products.

ROI in DoD Programs

Funding for many intelligence programs comes from the DoD, which requires all IT programs be managed as investments and not acquisitions. To achieve this goal and meet other government regulations and legislation such as the GPRA and the Information Technology Management and Results Act (ITMRA), the DoD has established performance measures in the IT investment process. Although profitability is not the primary goal of the DoD and other nonprofit organizations, there is pressure to ensure efficient use of taxpayers' money and deliver maximum value to citizens and communities.

Many issues are inherent in determining overall value and risks with ISR systems acquisitions. Technological complexities from the use of COT/GOTS systems, open architectures systems, evolving software standards, shortened acquisitions timelines, and funding instability all contribute to risks in Navy ISR systems. Although the DoD has instituted rigorous types of testing and evaluation (T&E) for all of its programs and projects to mitigate risk, metrics for IT systems have lacked the requisite depth for meaningful valuation. Crucial to successful T&E is the development of measurable key performance parameters (KPPs) and measures of effectiveness (MOEs) to provide accurate projections of system performance in a variety of operational environments.

Another issue in the DoD case is the translation of outputs into monetary benefits. Whereas in the commercial case, a price per unit is assigned to the outputs, there is no equivalent pricing mechanism in the DoD or nonprofit case. This presents a problem when conducting empirical financial analysis and in particular when seeking a baseline from which to formulate sound fiscal decisions. Valuation methodologies used by DoD for acquisitions must include a common framework for understanding, evaluating, and justifying the impact of government IT investments on the overall successful completion of the national security mission of the United States. KVA methodology is a viable valuation technique for that purpose.

Knowledge Value-Added Methodology

Knowledge Value-Added (KVA) was developed by Dr. Thomas Housel and Dr. Valery Kanevsky 15 years ago to estimate the value added by knowledge assets, both human and IT. It is based on the premise that businesses and other organizations produce outputs (e.g., products and services) through a series of processes and subprocesses, which change into raw inputs (i.e., labor into services, information into reports). Changes made on the inputs by organizational processes to produce outputs are the equivalent corresponding changes in entropy. Entropy is defined in the *American Heritage Dictionary* as a "measure of the degree of disorder [or change] in a closed system." In the business context, it can be used as a surrogate for the amount of changes that a process makes to inputs to produce the resulting outputs.[5]

Describing all process outputs in common units allows managers to assign revenues and costs to those units at any given point in time. With the resulting information, traditional accounting and financial performance and profitability metrics can be applied at the suborganizational level. KVA differs from other financial models in two important respects: It provides a method to analyze the metrics at a suborganizational level and allows for the allocation of cost and revenue across subprocesses for accounting purposes.

Knowledge value-added uses knowledge-based metaphor to operationalize the relationship between change in entropy and value added. The units of change induced by a process to produce an output are described in terms of the knowledge required to make the changes. More specifically, the time it takes the average learner "to acquire the procedural knowledge required to produce a process output provides a practical surrogate for the corresponding changes in entropy."[6]

The KVA, Monte Carlo simulation, and real options methodologies are applied to the USS *Readiness* in this case study to demonstrate how program managers can build metrics to conduct a financial analysis of each CCOP system at the process and subprocess levels. Managers and senior decision makers can thereby establish monetary values for traditionally intangible assets such as knowledge.

The USS *Readiness*

The goal of this case study is to assess the effectiveness and efficiency of CCOP systems in the Navy ISR mission. With KVA methodology, metrics are produced and the CCOP portfolio can be compared on existing and future programs. This section reviews how KVA is applied in two of the subprocesses in the CCOP program: Search/Collection Process (P4) and Format Data for Report Generation (P8).

The USS *Readiness* is a fictitious U.S. Navy warship outfitted to conduct ISR missions.[7] Along with the general manning, the ship has a contingent of IW operators performing intelligence collection processes utilizing CCOP systems. The ship is on a typical six-month deployment and receives daily tasking for ISR collection at national and tactical levels. Onboard the USS *Readiness* is an ISR crew of IW Officers: Division Officer, Division Leading Petty Officer, Signals Operators, and Comms Operators. Each IW officer performs certain processes in the ICP. After a request is received, the ISR crew produces a variety of reports that include raw intelligence reports, technical reports, analyst-to-analyst exchanges, and daily collection summaries. USS *Readiness* is outfitted with four CCOP systems (A, B, C, and D).

As shown in Table 14.4, CCOP systems may be used in a single subprocess or across multiple subprocesses along with the existing infrastructure available in each particular platform. Additionally, some systems such as CCOP A are highly complex and comprised multiple subsystems. With the help of KVA, the proxy revenues and costs are obtained and are shown in Table 14.5. Clearly, in the corporate setting, revenues and costs can be obtained quickly and easily, but KVA is required when applied to the public sector.

Table 14.6 lists the preliminary results where ROK is the return on knowledge (a productivity ratio), ROKA is the return on knowledge assets, a profitability ratio, and ROKI is the return on knowledge investment, the value equation.

The KVA provides the structured data required to perform various methods of risk analysis and performance projections such as real options analysis. This combination of KVA historical performance metrics, simulation, and real options analysis will enable the CCOP Program Office and the U.S. Navy to estimate and compare the future value added of different mixes

TABLE 14.4 USS *Readiness* CCOP Systems

	Subprocess Name	CCOP A	CCOP B	CCOP C	CCOP D
P1	Review request/tasking	X			
P2	Determine op/equip mix	X			
P3	Input search function/coverage plan	X			
P4	Search/collection process	X	X		
P5	Target data acquisition/capture	X	X		
P6	Target data processing	X	X	X	X
P7	Target data analysis	X		X	X
P8	Format data for report generation	X			
P9	QC report	X			
P10	Transmit report	X			

of human assets and systems as well as a range of new initiatives for the deployment and employment options of both.

Analyzing Real Options

A real options analysis was performed to determine the prospective value of three basic options over a 3-year period (Figure 14.14). The eight-step real options analysis process with KVA data was used to estimate the value of the options as seen earlier in this book.

The first option (A—Remote to Shore) was to use the various CCOP systems in a way that would allow all the data they generated to be viewed by a geographically remote center, the idea being that if all the intelligence collection processing could be done remotely in a consolidated center, fewer

TABLE 14.5 P4 and P8 Cost Allocation for CCOP C, D, and Fixed IT Infrastructure

Proxy Revenue Assigned to CCOP C Process K ($US)	Cost Assigned to CCOP C Process K ($US)	Proxy Revenue Assigned to CCOP D Process K ($US)	Cost Assigned to CCOP D Process K ($US)	Proxy Revenue Assigned to Fixed Infras Process K ($US)	Cost Assigned to Fixed Infras Process K ($US)
$	$			$ 28,156	$ 10,250
$	$			$ 13,868	$ 10,250
$58,253	$12,000	$19,906	$63,462	$241,667	$102,500

TABLE 14.6 P4 and P8 KVA Metrics

		KVA Metrics for Total K			
	Subprocess Name	**ROK as Ratio**	**ROK (%)**	**ROKA (%)**	**ROKI (%)**
P4	Search/collection	3.39	339.01	70.50	239.01
P8	Format data for report generation	0.80	79.63	–25.59	–20.37
Metrics for aggregated		14.10	1410.20	157.31	410.20

intelligence personnel would be required on ships. The idea of remoting capabilities to a consolidated center is a popular movement in the military to cut costs and provide more shore-based operations to support warfighting capabilities. This is akin to the consolidation of service operations in businesses—for example, in larger, but fewer, call centers.

The second option (B—Direct Support) focused on how the CCOP's equipment and operators could be moved from ship to ship. When a ship came into port for maintenance, repair, or modernization, the idea was to move the CCOP equipment and operators to ships that were about to be deployed. This way, fewer sets of CCOP equipment and operators would be needed to service the intelligence gathering needs of the fleet.

The third option (C—Permanent SSES) basically kept the CCOP systems and operators assigned to given ships at all times. This approach required more operators and CCOP systems raising the potential costs but providing more control of the intelligence capability by the ships and fleet commanders.

The results of the analysis (Figure 14.15) indicated that the highest value was for option C. The result ran contrary to the expected cost savings of options A and B. However, because KVA provided a monetized numerator in the form of surrogate revenue, it was possible to see the effects of greater outputs-revenue for option C. Option C is the preferred option of the commanders of the fleet and ships because it affords greater control of the intelligence assets for their specific operations. So, intuitively, these commanders favored option C, but prior to the real options with KVA data analysis, they had no relatively objective way to support their intuitions.

It is possible that with time and experience, the remoting option would provide greater benefits-revenue per cost than data collection techniques because remoting provides more robust operations from ship platforms. But, the current bandwidth limitations of the naval operating environment mitigate against remoting systems that have high bandwidth requirements.

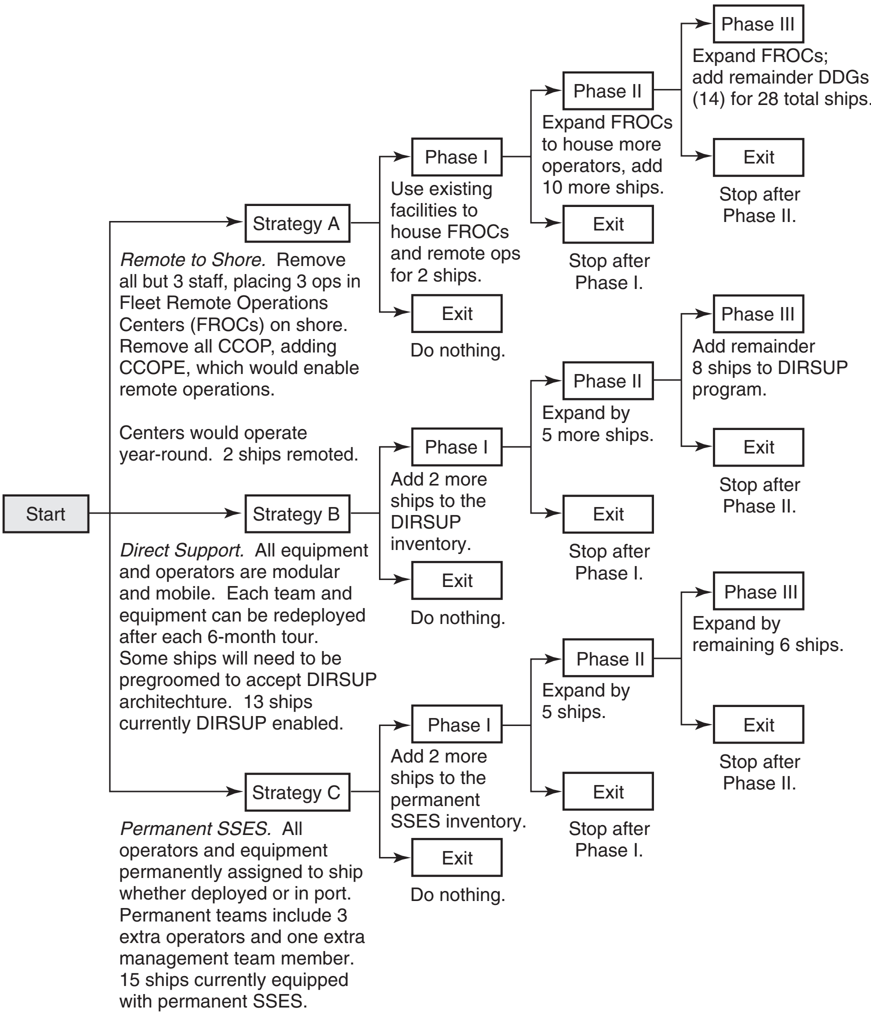


FIGURE 14.14 Staging three path-dependent real options strategies for CCOPs.

The CCOP's program office has asked for further analysis using the KVA and real options methodologies. Software that applies KVA, simulation, and real options analysis are routinely in the process of being deployed with a naval strike group to enable ongoing monitoring of the performance of the data collection process and its supporting CCOP systems. The next step will be to include the use of this software to enable the commanders and program executives to make projections about the best options for

Summary Results	Strategy A	Strategy B	Strategy C
PV Option Cost (Year 1)	$348,533	$1,595,697	$1,613,029
PV Option Cost (Year 2)	$4,224,487	$3,043,358	$4,494,950
PV Option Cost (Year 3)	$3,688,994	$10,105,987	$8,806,643
PV Revenues	$24,416,017	$33,909,554	$48,420,096
PV Operating Costs	$16,220,188	$16,765,513	$9,951,833
PV Net Benefit	$8,195,829	$17,144,041	$28,868,264
PV Cost to Purchase Option	$425,000	$169,426	$72,611
Maturity in Years	3.00	3.00	3.00
Average Risk-Free Rate	3.54%	3.54%	3.54%
Dividend Opportunity Cost	0.00%	0.00%	0.00%
Volatility	26.49%	29.44%	15.04%
Total Strategic Value with Options	$1,386,355	$4,466,540	$15,231,813

FIGURE 14.15 Summary real options analysis results.

deploying the CCOP systems to support the intelligence needs of the naval commanders and other intelligence gathering and analysis agencies in the federal government.

CASE STUDY: MANUFACTURING AND SALES IN THE AUTOMOTIVE AFTERMARKET

This case study was written by Andy Roff and Larry Blair with modeling assistance from the author. Both Andy and Larry are executives from the automotive aftermarket who have owned and managed several businesses. They each have 30 plus years of experience, specifically in the provision of information systems for the shared benefit of both suppliers and distributors. They can be contacted at larblair2@aol.com.

Background and History of the Automotive Aftermarket

The automotive aftermarket (AAM) started soon after the first horseless carriage made its way on to the world's roads more than a century ago. It happened perhaps within a couple of days when the original dog-clutch gave in to the abuse of its erstwhile horse-driving operator! And thus, the AAM was born the moment the first screw needed replacement. Over time, as makes and models of automobiles multiplied, so did the manufacturers of parts to repair and keep them running—some commissioned by the auto makers—with various manufacturing "pattern" parts of varying quality and durability. As of the time of writing, the world APA is approximately $800 billion per year, with a 3 percent expected growth rate going forward.

With so many different parts and suppliers, there was a need for reference books to identify the correct item, so giving birth to the parts catalog. Nothing much changed until the introduction of the microfiche in the 1960s, and that was used almost exclusively by the car makers' service and parts network. In skilled hands, this quasi-electronic database brought efficiency and speed to the parts sales and automotive repair processes. However, the wealth and complexity of the largely graphic-based content made its adoption by the competitive aftermarket nonviable.

Instead, the ubiquitous personal computer became a natural tool for the advent of electronic cataloguing in the early 1980s. A major hurdle that still had to be overcome was that the various proprietary systems demanded a high level of specific data formatting, which was an extremely costly exercise to undertake and conflicted with the existing print-oriented legacy practices of the catalog authors. Also, there was little point in the major aftermarket suppliers each devising and installing their own e-catalog versions when each one would demand a separate hardware platform to run on. Worse, these platforms could not integrate with the computerized point-of-sale (PoS) tills that were introduced in the late 1970s.

This demand vacuum was just too big, both conceptually and given the constraints of the available technology. The first European attempt concentrated on providing a standalone terminal, bringing together parts from multiple providers in a "bookcase" format. There was a common drill-down of available car makes and models with access to information compiled by each aftermarket supplier and designed for dissemination by trade associations. This system originated in the Netherlands and was also licensed for use in the United Kingdom during the early 1990s.

Although eventually a commercial failure, this system's introduction forced the parts suppliers and manufacturers to focus on e-data provision and begin the shift from a print-centric catalog-building mentality. This shift was reinforced by the ambitions of national parts distribution chains to provide "tied" e-catalogs and the leading PoS providers to add e-cataloguing capabilities to their terminals. Both initiatives increased the demands for e-data from suppliers and manufacturers.

In the United States, a national PoS provider decided on a massive investment in 1984, leading to the introduction of a dedicated, integrated e-catalog in 1985, followed by a European version 5 years later.

So the manufacturers' primary "shop window" took the form of these third-party e-cataloguing systems. They were forced to become less possessive about their data, had less control of timeliness and accuracy in the way it was presented to the marketplace, and had to provide multiple versions of the e-catalogs for the various national chains and third-party providers. In some cases, they were even obliged to pay to have it placed on display.

The Issues Facing the Industry

There is a silver lining to this particular cloud. Given that the formats of the data are now becoming increasingly common and indexed to an industry standard—in the United States, industry-sponsored lists are available—data will become increasingly consistent with a faster time to market. Now that technology is more advanced, graphics and illustrations on the part's characteristics, its location on the car, fitting tips, and other key information can all be linked into such a catalog. All of these improvements will help to increase the quality of the buying experience and enable individual manufacturers the ability to distinguish their offerings.

And there's more. Manufacturers' products can be accurately linked to a list of cars and that list of cars can be linked to state-provided car population statistics. Now production and distribution strategies can be subjected to risk analysis, simulation, forecasting, optimization, and real options analysis. When all the possible components impacting on a decision to manufacture or source the supply of a given replacement part are taken into consideration, it shows just how fragile and error-prone traditional decision-making methods must be.

The Analytical Complexity

An example case study is based on Casky Automotive Electrics, Inc., a theoretical private company specializing in the design and manufacturing of automotive products in support of the original equipment manufacturing (OEM) sector of the automotive industry. Casky's specialty is rotating electrics, commonly known as *alternators* and *starters*. The company has close ties with both Ford and General Motors (GM), and these firms have provided Casky with a basis for growing their business in both North America and Europe. As a development partner to two of the world's largest automotive manufacturers, Casky has supported the development programs of both manufacturers with engineering expertise that has led to manufacturing contracts for starter motors for some of the most recognized car models on the road. Those relationships have also led to contracts for some of the newest hybrid models in which fuel efficiency is maximized. These models place an even greater burden on the starter motor and therefore increase its cost and complexity. Casky has won a contract for the starter motor for the hypothetical new Phalynx hybrid minivan that was introduced by GM in 2005. GM has placed orders for the units that will be fitted to the cars during their assembly. However, Casky has also won the contract for the service support of the dealer network for replacement of starters as required by service demands. The automotive manufacturing industry is one of the largest (considering both economic value and employment) in those countries having a high vehicle registration. Certainly North America and Europe account for most of the vehicle registrations in the world and highest per capita ratios of car registrations in comparison to the general population.

Initial total sales of the Phalanx are predicted at 100,000 per year, rising to 150,000 in year 2 and reducing to 100,000 in year 3. Sales predictions for similar models in the past have been accurate to ±5 percent. The vehicle will be manufactured in mainland Europe and primarily marketed in Europe and North America. The eventual population of the model will vary across the European and North American states but will aggregate 55 percent and 45 percent in the two markets. Vehicle population statistics will be available annually from various external suppliers. A face-lift, as opposed to an all-new, model will be marketed in year 4, and sales are predicted to recover to 150,000 before declining steadily to 75,000 in year 5 prior to an all-new model launch. The total predicted model population will therefore be 575,000 with an annual scrap rate—attrition through either insurance total loss or being uneconomical to repair—of 2 percent compounded annually. There are two gasoline and one diesel versions with a prediction of equal demand for all three engine variations across the model range, with these engines serving both the original and face-lift versions, but not the all-new model.

Caskey is chosen to provide the starting motor for all three engines. They supply only new, as opposed to reconditioned, units both to GM and the automotive aftermarket (AAM). Each starting motor unit is different, having been specifically designed for a specific model, and has different wear characteristics with a minimum time before failure (MTBF) of 100,000 miles for the smaller gas engine, 85,000 miles for the larger gas engine, and 100,000 miles for the diesel version. The average annual user mileage is predicted at 12,000 for the smaller gas engine and 15,000 for both the larger gas engine and the diesel. There is a 2-year warranty on the units sold in mainland Europe and 3 years in the United Kingdom, Ireland, and North America.

There is a demand from GM of sufficient stock on hand for 1 week's production with a zero failure rate at fitting. The unexpected failure rate (that is, before MTBF and therefore resulting in a warranty claim) is 1:10,000. GM's retail service network has 250 outlets in Europe and 150 in North America, and each must hold at least two of each unit at the model launch. Caskey has three European and two North American distribution warehouses that service both GM's retail network and the AAM through both independent and chain parts retailers. The margin is the least on sales to the GM, +20 percent to the national chains, and +25 percent to the independents.

Caskey expects to supply 90 percent of units sold through GM's service network outside of warranty claims, but competes from the 4th year onward with other new unit manufacturers and in the 5th year onward with unit reconditioners. There is a single European new unit manufacturer with a distribution network in North America that introduces a modified starting motor, which also fits another model with an existing and out-of-warranty model with a similar population and engine mix. This new unit manufacturer expects to gain an initial 10 percent of the market for the new model, rising by 2 percent compound and 50 percent of the additional model where it was one of two vehicle parts manufacturers (VPMs) selected for the original equipment. Two unit reconditioners enter the market in North America and three in Europe, each expecting a 5 percent share of the market and each distributing only to the AAM. The reconditioners' ability to service the market is directly related to the return of worn units which in the first year of their operation (year 5 of production) is 100 percent from GM, reducing in subsequent years as more units from the other new unit VPM wear out. The MTBF for the reconditioned units is only 66 percent of that for the all-new units.

The Analytical Framework Applying Risk Analysis, Simulation, Forecasting, and Optimization

Setting up and solving the problem is not a trivial task, requiring facility with Risk Simulator's Monte Carlo simulation, forecasting, and optimization routines. Figure 14.16 illustrates a forecast model of the automobile demand based on the assumptions listed previously. Minimum, maximum, and most likely value ranges are also listed and each of the period's demand values is simulated (Figure 14.17); that is, the European and U.S. demands for each quarter are simulated such that the expected values of each year are in line with the foregoing assumptions of 100, 150, 100, 150, and 75 thousand vehicles, respectively.

Figure 14.18 illustrates the modeling of the additional requirements and restrictions of the demand forecasts, such as failure rates of the parts, scrap rates of the automobile model, and average miles driven per year. Note that all the highlighted cells in Figures 14.16 and 14.18 are simulation assumptions and each value is simulated thousands of times in the model. Next, an optimization model is developed based on these uncertainties in demand levels, as shown in Figure 14.19. In this model, we see that the decision variables are the quantity to manufacture; that is, to find the optimal quantity to manufacture given the uncertainty-based forecasted demand levels. Price per unit, failure rates, and average driving distance per year for a vehicle are all accounted for in the model. The analysis provides the optimal quantity to manufacture such that the total net profits are maximized, subject to excess costs of surplus and shortages in quantity on hand.

						Possible Ranges for Actual Auto Demand					
Period	Europe	USA	Total	Running Total	Annual Totals	Minimum Europe	Most Likely Europe	Maximum Europe	Minimum USA	Most Likely USA	Maximum USA
Year 1 Q1	11,001	8,996	20,000	20,000		10,451	11,001	11,551	8,546	8,996	9,445
Year 1 Q2	13,767	11,242	25,000	45,000		13,079	13,767	14,455	10,680	11,242	11,804
Year 1 Q3	19,266	15,763	35,000	80,000		18,303	19,266	20,230	14,974	15,763	16,551
Year 1 Q4	10,998	8,999	20,000	100,000	100,000	10,448	10,998	11,547	8,549	8,999	9,449
Year 2 Q1	16,497	13,504	30,000	130,000		15,672	16,497	17,322	12,828	13,504	14,179
Year 2 Q2	20,615	13,504	37,500	167,500		19,584	20,615	21,646	16,057	16,902	17,747
Year 2 Q3	28,892	23,629	52,500	220,000		27,447	28,892	30,336	22,448	23,629	24,810
Year 2 Q4	16,499	13,503	30,000	250,000	150,000	15,674	16,499	17,324	12,828	13,503	14,178
Year 3 Q1	10,997	8,995	20,000	270,000		10,447	10,997	11,547	8,546	8,995	9,445
Year 3 Q2	13,746	11,246	25,000	295,000		13,059	13,746	14,434	10,684	11,246	11,809
Year 3 Q3	19,253	15,737	35,000	330,000		18,290	19,253	20,216	14,951	15,737	16,524
Year 3 Q4	11,004	9,007	20,000	350,000	100,000	10,453	11,004	11,554	8,557	9,007	9,458
Year 4 Q1	16,500	13,498	30,000	380,000		15,675	16,500	17,325	12,823	13,498	14,173
Year 4 Q2	20,619	16,890	37,500	417,500		19,588	20,619	21,650	16,046	16,890	17,735
Year 4 Q3	28,882	23,637	52,500	470,000		27,438	28,882	30,326	22,455	23,637	24,819
Year 4 Q4	16,487	13,504	30,000	500,000	150,000	15,663	16,487	17,311	12,828	13,504	14,179
Year 5 Q1	8,246	6,757	15,000	515,000		7,834	8,246	8,658	6,419	6,757	7,094
Year 5 Q2	10,307	8,433	18,750	533,750		9,792	10,307	10,823	8,012	8,433	8,855
Year 5 Q3	14,434	11,802	26,250	560,000		13,712	14,434	15,156	11,212	11,802	12,392
Year 5 Q4	8,255	6,746	15,000	575,000	75,000	7,843	8,255	8,668	6,409	6,746	7,083
Grand Total	316,266	258,790	575,000			300,452	316,266	332,079	245,850	258,790	271,729

FIGURE 14.16 Automobile demand forecast.

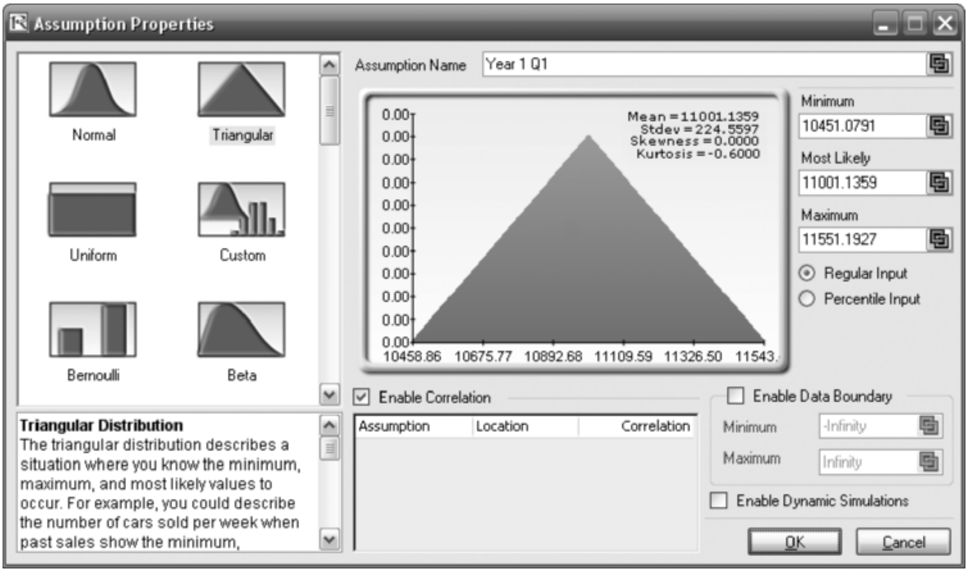


FIGURE 14.17 Monte Carlo simulation of demand forecast.

Initial Projected Scrap Rate	1.99%	(Range is from 1.50% to 2.5% per year)	
Postwarranty Scrap Rate	10.82%	(Range is from 8% to 15% per year)	
Projected Miles Driven Per Year		12,000	(Range is from 10,000–14,000 for small petrol engines)
		15,000	(Range is from 13,000–17,000 for large petrol engines)
		14,989	(Range is from 13,000–17,000 for diesel engines)
Average Warranty		100,000 (miles)	
Prewarranty Failure Rate	0.01%	(Range is from 0.01% to 0.02% per week)	
Postwarranty Failure Rate	0.15%	(Range is from 0.05% to 0.20% per week)	

FIGURE 14.18 Additional requirements.

For instance, say we have a marginal holding or carrying cost of $1.00 for each surplus unit manufactured versus a cost of $1.20 marginal excess net losses in sales if there is a shortage in manufactured parts with respect to sales demand. In addition, at least 800 units must be available within the first 6 months to cover the two-unit minimum per outlet for the 400 outlets worldwide. Finally, the manufactured output cannot exceed 1.50 times the forecasted values per year, to prevent any glut in the market. Monte Carlo simulation and forecasting methodologies were applied as well as dynamic

Expected Period	Auto Parts Demand Forecast Europe	USA	Total	Miles Driven		Quantity to Manufacture	Required Min	Required Max	Shortage or Surplus	Marginal Cost	Total Sales	Price Unit	Min Price	Max Price	Stochastic Sales
Year 1 Q1	83	68	152	3,747		400	200	500	248	$ (248.21)	$ 11,384.55	$ 74.88	$ 50.00	$100.00	$ 11,291.05
Year 1 Q2	188	154	342	7,495		400	200	500	58	$ (58.46)	$ 25,615.24	$ 74.78	$ 50.00	$100.00	$ 25,539.16
Year 1 Q3	334	273	607	11,242		224	200	500	–383	$ (459.81)	$ 18,800.00	$ 75.06	$ 50.00	$100.00	$ 16,812.41
Year 1 Q4	417	342	759	14,989		225	200	500	–534	$ (640.76)	$ 16,875.00	$ 74.65	$ 50.00	$100.00	$ 16,795.51
Year 2 Q1	543	444	987	18,736		977	300	2,000	–10	$ (11.59)	$ 109,912.50	$ 112.31	$ 75.00	$150.00	$109,729.00
Year 2 Q2	699	527	1,271	22,484		978	300	2,000	–293	$ (351.93)	$ 110,025.00	$ 112.00	$ 75.00	$150.00	$109,540.33
Year 2 Q3	918	751	1,670	26,231		978	300	2,000	–692	$ (830.08)	$ 110,025.00	$ 112.55	$ 75.00	$150.00	$110,078.79
Year 2 Q4	1,044	854	1,897	29,978		978	300	2,000	–919	$ (1,103.31)	$ 110,025.00	$ 111.30	$ 75.00	$150.00	$108,853.36
Year 3 Q1	1,127	922	2,049	33,725		1,586	500	4,000	–463	$ (555.86)	$ 237,900.00	$ 148.90	$100.00	$200.00	$236,163.21
Year 3 Q2	1,231	1,008	2,239	37,473		1,586	500	4,000	–653	$ (783.55)	$ 237,900.00	$ 151.04	$100.00	$200.00	$239,546.39
Year 3 Q3	1,378	1,127	2,505	41,220		1,586	500	4,000	–919	$ (1,102.32)	$ 237,900.00	$ 147.99	$100.00	$200.00	$234,717.81
Year 3 Q4	1,461	1,195	2,656	44,967		1,586	500	4,000	–1,069	$ (1,283.27)	$ 238,050.00	$ 148.44	$100.00	$200.00	$235,580.13
Year 4 Q1	1,586	1,298	2,884	48,715		2,251	500	4,000	–633	$ (759.70)	$ 337,650.00	$ 149.12	$100.00	$200.00	$335,671.99
Year 4 Q2	1,743	1,426	3,169	52,462		2,252	500	4,000	–917	$ (1,100.04)	$ 337,800.00	$ 152.04	$100.00	$200.00	$342,396.00
Year 4 Q3	1,961	1,605	3,567	56,209		2,252	500	4,000	–1,315	$ (1,578.19)	$ 337,800.00	$ 150.20	$100.00	$200.00	$338,241.96
Year 4 Q4	2,087	1,708	3,795	59,956		2,252	500	4,000	–1,543	$ (1,851.42)	$ 337,800.00	$ 150.49	$100.00	$200.00	$338,897.62
Year 5 Q1	2,150	1,759	3,909	63,704		2,782	500	4,000	–1,127	$ (1,352.04)	$ 417,300.00	$ 151.17	$100.00	$200.00	$420,552.35
Year 5 Q2	2,228	1,824	4,051	67,451		2,782	500	4,000	–1,269	$ (1,522.80)	$ 417,300.00	$ 150.71	$100.00	$200.00	$419,276.61
Year 5 Q3	2,338	1,913	4,250	71,198		2,782	500	4,000	–1,468	$ (1,761.88)	$ 417,300.00	$ 150.03	$100.00	$200.00	$417,384.52
Year 5 Q4	2,400	1,964	4,364	74,946		2,782	500	4,000	–1,582	$ (1,898.49)	$ 417,300.00	$ 148.95	$100.00	$200.00	$414,391.63
Year 6 Q1	2,353	1,925	4,277	78,693		2,786	500	4,000	–1,491	$ (1,789.63)	$ 417,900.00	$ 151.00	$100.00	$200.00	$420,674.40
Year 6 Q2	2,306	1,887	4,192	82,440		2,786	500	4,000	–1,406	$ (1,687.63)	$ 417,900.00	$ 149.22	$100.00	$200.00	$415,723.94
Year 6 Q3	2,260	1,849	4,109	86,187		2,787	500	4,000	–1,322	$ (1,546.46)	$ 418,050.00	$ 148.80	$100.00	$200.00	$414,703.62
Year 6 Q4	2,215	1,812	4,027	89,935		2,787	500	4,000	–1,240	$ (1,488.47)	$ 418,050.00	$ 148.93	$100.00	$200.00	$415,055.05
Year 7 Q1	2,171	1,776	3,947	93,682		3,947	500	4,000	0	$ (0.43)	$ 592,050.00	$ 149.16	$100.00	$200.00	$688,741.26
Year 7 Q2	2,128	1,71	3,869	97,429		3,869	500	4,000	0	$ (0.08)	$ 580,338.17	$ 150.50	$100.00	$200.00	$582,285.22
Year 7 Q3	9,871	8,076	17,948	101,176	*Warranty Expires*	4,000	500	4,000	–13,948	$ (16,737.06)	$ 600,000.00	$ 248.15	$100.00	$200.00	$593,796.91
Year 7 Q4	8,804	7,203	16,006	104,924		4,000	500	4,000	–12,006	$ (14,407.69)	$ 600,000.00	$ 150.91	$100.00	$200.00	$603,657.11
Year 8 Q1	7,851	6,424	14,275	108,671		4,000	500	4,000	–10,275	$ (12,330.26)	$ 600,000.00	$ 151.07	$100.00	$200.00	$604,279.89
Year 8 Q2	7,002	5,729	12,731	112,418		4,000	500	4,000	–8,731	$ (10,477.52)	$ 600,000.00	$ 150.93	$100.00	$200.00	$603,708.20
Year 8 Q3	6,245	5,109	11,354	116,166		4,000	500	4,000	–7,453	$ (8,825.16)	$ 600,000.00	$ 149.51	$100.00	$200.00	$598,020.37
Year 8 Q4	5,569	4,557	10,126	119,913		4,000	500	4,000	–6,126	$ (7,351.51)	$ 600,000.00	$ 151.25	$100.00	$200.00	$604,988.39
Year 9 Q1	4,967	4,064	9,031	123,660		4,000	500	4,000	–5,031	$ (6,037.25)	$ 600,000.00	$ 150.33	$100.00	$200.00	$601,315.42
Year 9 Q2	4,430	3,624	8,054	127,407		4,000	500	4,000	–4,054	$ (4,865.13)	$ 600,000.00	$ 150.79	$100.00	$200.00	$603,151.05
Year 9 Q3	3,951	3,232	7,183	131,155		4,000	500	4,000	–3,183	$ (3,819.79)	$ 600,000.00	$ 150.17	$100.00	$200.00	$600,671.48
Year 9 Q4	3,523	2,883	6,406	134,902		4,000	500	4,000	–2,406	$ (2,887.50)	$ 600,000.00	$ 149.10	$100.00	$200.00	$596,392.93
Year 10 Q1	3,142	2,571	5,713	138,649		3,858	500	4,000	–1,856	$ (2,226.45)	$ 578,700.00	$ 150.14	$100.00	$200.00	$579,252.36
Year 10 Q2	2,802	2,293	5,095	142,396		3,858	500	4,000	–1,237	$ (1,484.93)	$ 578,700.00	$ 149.96	$100.00	$200.00	$578,548.04
Year 10 Q3	2,499	2,045	4,544	146,144		3,859	500	4,000	–685	$ (822.40)	$ 578,850.00	$ 150.89	$100.00	$200.00	$582,296.17
Year 10 Q4	2,229	1,824	4,053	149,891		3,859	500	4,000	–194	$ (232.61)	$ 578,850.00	$ 148.00	$100.00	$200.00	$571,137.56
Year 11 Q1	1,988	1,627	3,614	153,638		2,531	500	4,000	–1,083	$ (1,300.20)	$ 379,650.00	$ 149.92	$100.00	$200.00	$379,435.92
Year 11 Q2	1,773	1,451	3,224	157.496		2,531	500	4,000	–693	$ (831.08)	$ 379,650.00	$ 150.23	$100.00	$200.00	$380,226.69
Year 11 Q3	1,581	1,294	1,875	161,133		2,532	500	4,000	–343	$ (411.50)	$ 379,800.00	$ 150.56	$100.00	$200.00	$381,210.00
Year 11 Q4	1,410	1,154	2,564	164,880		2,532	500	4,000	–32	$ (38.37)	$ 379,800.00	$ 148.24	$100.00	$200.00	$375,336.46
Year 12 Q1	1,258	1,029	2,287	168,627		2,000	300	2,000	–287	$ (344.00)	$ 300,000.00	$ 150.02	$100.00	$200.00	$300,038.57
Year 12 Q2	1,122	918	2,039	172,375		1,661	300	2,000	–378	$ (454.02)	$ 249,150.00	$ 150.70	$100.00	$200.00	$250,312.34
Year 12 Q3	1,000	818	1,819	176,122		1,661	300	2,000	–158	$ (189.34)	$ 249,150.00	$ 149.70	$100.00	$200.00	$248,656.93
Year 12 Q4	892	730	1,622	179,869		1,322	300	2,000	–300	$ (360.08)	$ 198,300.00	$ 150.56	$100.00	$200.00	$199,034.29
Year 13 Q1	796	651	1,447	183,617		500	200	500	–947	$ (1,135.96)	$ 75,000.00	$ 151.45	$100.00	$200.00	$ 75,726.91
Year 13 Q2	710	581	1,290	187,364		500	200	500	–790	$ (948.20)	$ 75,000.00	$ 150.21	$100.00	$200.00	$ 75,106.59
Year 13 Q3	633	518	1,151	191,111		500	200	500	–651	$ (780.76)	$ 75,000.00	$ 149.66	$100.00	$200.00	$ 74,829.28
Year 13 Q4	564	462	1,026	194,858		500	200	500	–526	$ (631.42)	$ 75,000.00	$ 150.37	$100.00	$200.00	$ 75,186.06

FIGURE 14.19 Optimization model.

optimization techniques. The actual part quantities that should be manufactured that maximize net profits, minimize excess losses, and are all the while subject to the relevant minimum and maximum manufactured parts are illustrated in Figure 14.19 and charted in Figure 14.20. As can be seen in the chart, it is optimal to start with a small quantity initially when the Phalanx is introduced, and gradually but with a stepwise progression, increase the number of parts as the car gets older. The quantity peaks between years 7 and 10 when warranties expire and when the parts are most needed, and then gradually decreases over time as the cars are decommissioned, sold, or scrapped.

Using these advanced risk analysis techniques, we are able to predict the optimal manufacturing output and the life cycle of a specific part based on historical data and simulating thousands of potential outcomes and scenarios in an optimization model. In fact, we can take this one step further and on completion of the optimization analysis, reapply simulation and obtain the probabilities of the net revenues for this particular part, as seen in Figures 14.21, 14.22, and 14.23.

However, we require two parts per outlet for 250 outlets
in Europe and 150 outlets in the United States.
So, the first period requires 800 units.

Assumed Cost of a Surplus unit: $1.00 (Additional carrying cost losses per unit)
Assumed Cost of a Shortage unit: $1.20 (Additional sales loss per unit)

Constraints:
1. First 6 months must be at least 800 units: 800
2. Each year we cannot manufacture more than 1.5 times the forecasted demands:

Year	Manufactured	Limit
1	1,249	1,859
2	3,911	5,825
3	6,345	9,449
4	9,007	13,415
5	11,128	16,574
6	11,146	16,606
7	15,816	41,770
8	16,000	48,487
9	16,000	30,675
10	15,434	19,406
11	10,126	12,277
12	6,644	7,767
13	2,000	4,914

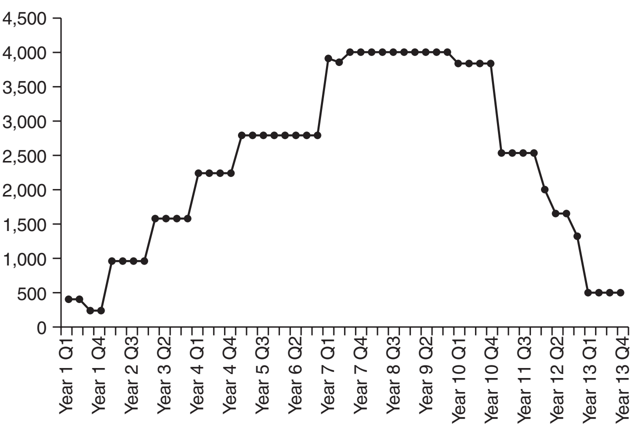


FIGURE 14.20 Optimal quantity and manufacturing constraints.

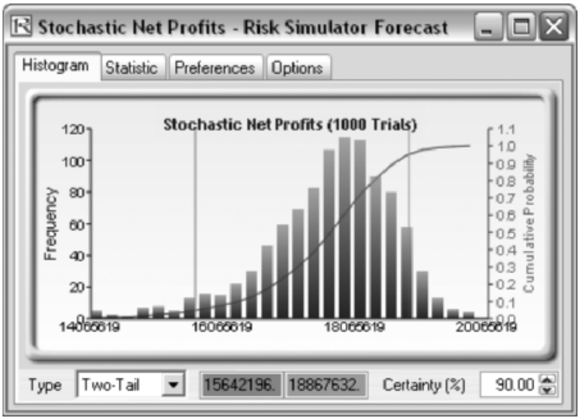


FIGURE 14.21 The 90 percent confidence interval.

Statistics	Result
Number of Trials	1000
Mean	1.754064E+007
Median	1.768675E+007
Standard Deviation	9.702058E+005
Variance	9.412994E+011
Average Deviation	7.463943E+005
Maximum	1.977128E+007
Minimum	1.391331E+007
Range	5.857964E+006
Skewness	-0.8801
Kurtosis	1.1070
25% Percentile	1.701109E+007
75% Percentile	1.821140E+007
Percentage Error Precision at 95% Confidence	0.3428%

FIGURE 14.22 The simulated statistics.

Figure 14.21 shows that the 90 percent confidence interval of the net profits for this particular part is between \$15.64 and \$18.87 million over its lifetime. In fact, the expected value or mean net profit is \$17.54 million (Figure 14.22). Finally, using the simulated results, we can compare the profitability of one part versus another. For instance, suppose we have an alternative part that the company is deciding on manufacturing and the expected net profit payoff is \$15.0 million. We can determine that by manufacturing the current parts, there is a 91.20 percent probability that this current part's net profits will exceed the alternative business line.

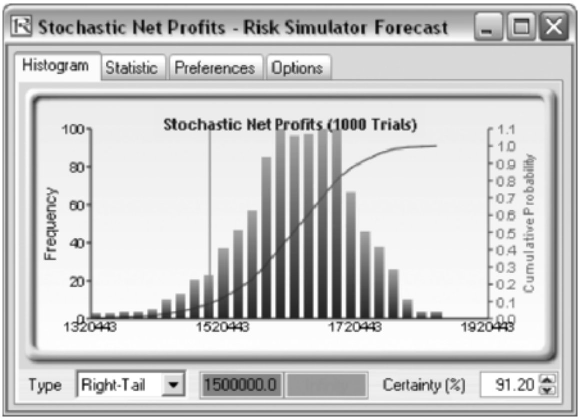


FIGURE 14.23 Sample breakeven points.

In contrast, had optimization and simulation risk analysis not been performed, the results would have been a highly suboptimal set of results. For instance, based on the required minimum and maximum production required in each period, say we manufacture at the average of the forecasted values; the total net profits would have been \$13.43 million or manufacturing at the required minimum required values returns \$0.71 million in net profits. Therefore, given such huge swings in values, running optimization guarantees the maximum possible net profits of \$17.54 million subject to the uncertainties and risks inherent in the demand forecasts.

To conclude, Monte Carlo simulation, forecasting, and optimization are crucial in determining the risk elements and uncertainties of pricing and demand levels. In addition, the analysis provides a set of valid optimal quantities to manufacture given these uncertainty demand levels, all the while considering the risk of the business line. Thus, using risk analysis, decision makers can not only decide which business lines or parts to manufacture, but how much to manufacture, when to manufacture them, and if required, to decide the optimal price points to sell the parts, maximize profits, and minimize any losses and risks.

CASE STUDY: THE BOEING COMPANY'S STRATEGIC ANALYSIS OF THE GLOBAL EARTH OBSERVATION SYSTEM OF SYSTEMS

This case study was written by Ken Cobleigh, Dan Compton, and Bob Wiebe, from The Boeing Company in Seattle, Washington, with assistance from the author. This is an actual consulting project performed by Ken, *Dan, Bob and the author on the GEOSS system. Although the facts are correct at the time of writing, the analysis has been significantly simplified for the purposes of this case study.*

A Background on the Global Earth Observation System of Systems

On February 16, 2005, 61 countries agreed to a plan that, over the next 10 years, humanity will revolutionize its understanding of the earth and how it works. Agreement for a 10-year implementation plan for a Global Earth Observation System of Systems, known as GEOSS, was reached by member countries of the Group on Earth Observations at the Third Observation Summit held in Brussels. Nearly 40 international organizations also support the emerging global network. The number of participating countries has nearly doubled, and interest has accelerated since the December 2004 tsunami devastated parts of Asia and Africa. In the coming months, more countries and global organizations are expected to join the historic initiative. The GEOSS project will help all nations involved produce and manage their information in a way that benefits the environment and humanity by taking the pulse of the planet. The beneficiaries are divided into nine major categories, as depicted in Figure 14.24.

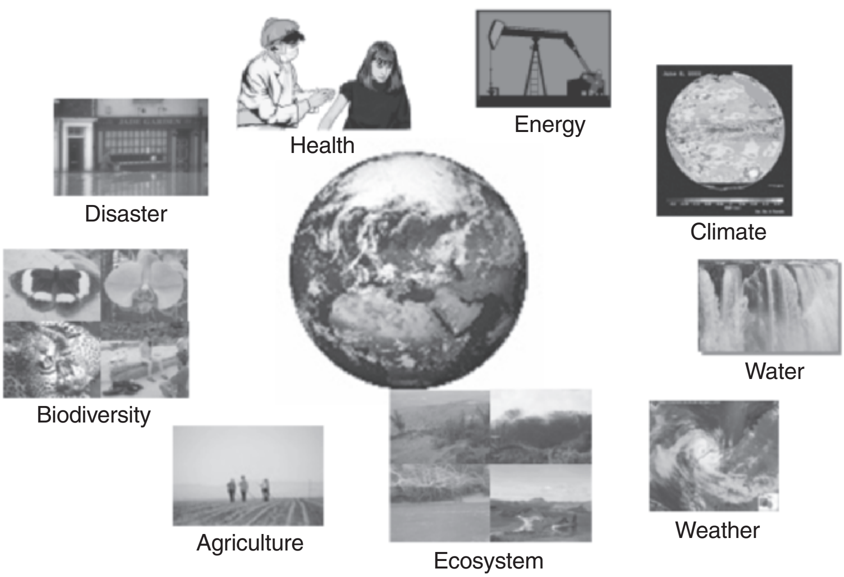


FIGURE 14.24 Societal benefits from earth observations.

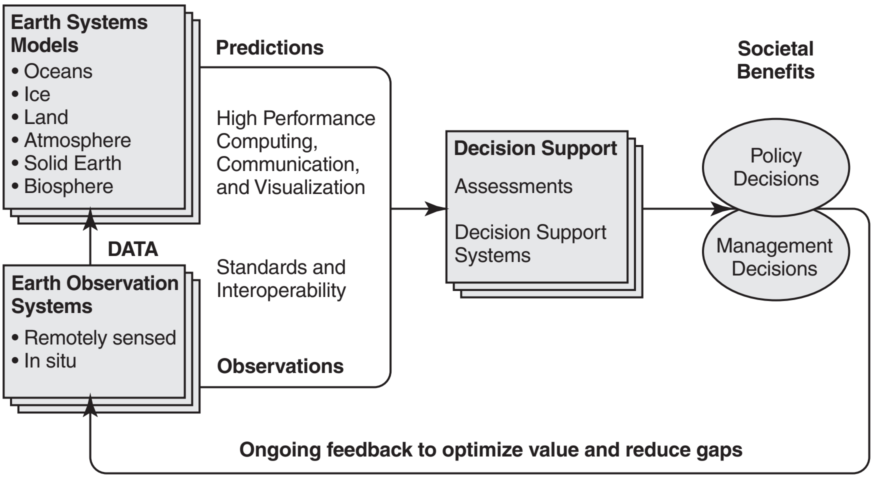


FIGURE 14.25 GEOSS dynamic decision process.

The data can come from satellites, airplanes, balloons, ships, radars, river gauges, ground weather stations, buoys, and field data collected on recorders as well as data collected with pencil and paper. An end-to-end architecture was derived from the basic needs of the system and defined into three groups: the observations systems and other domain data; the GEOSS information architecture; and the user communities. This is shown in Figures 14.25 and 14.26. Some of the applications can be as basic as measuring an ecosystem's biodiversity of animal life to measuring, capturing, analyzing, and better predicting natural disasters like tsunamis, earthquakes, hurricanes, and so forth, providing a global early warning system, saving lives in the process.

Currently, several issues must be overcome in order to allow a long-term high-level vision such as the GEOSS to become a reality. An assessment was made with the technical GEOSS community, which comprised several subcommittees; the one the authors consulted with was the architecture subcommittee. The major issues are summarized in the following list:

- Capability of supporting multiple data formats and exchanging between formats.
- Agree on a new standard format for raw and processed data for new systems.
- Provide information assurance (knowing the data will arrive uncorrupted).
- Provide data, information security, and controlled access (country restrictions, classified data, and so forth).

- Assure easy use of data and information including training, data mining, and other usability tools.
- Enable the creation and use of decision support tools.
- Allow data and knowledge products (higher-level processed products through the use of multiple sensor fusing).
- Assure easy global access.
- Allow data and knowledge products (higher level processed products through the use of multiple sensor and nonsensor fusing).
- Provide high throughput end to end.
- Support nonelectronic transfer of data and information.
- Provide low latency.

As can be seen, many of these high level issues are going to be politically and economically charged. For instance, is the economic benefit decided by the country's gross domestic product or wealth, or by the countries that are in the most need of the benefits? Clearly, a lot of discussions and negotiations need to occur before such a system can be realized. And most likely, it will happen in stages.

One current problem is that many systems are built as stand-alone or stovepipe systems. Their data does not easily register, correlate, or fuse with

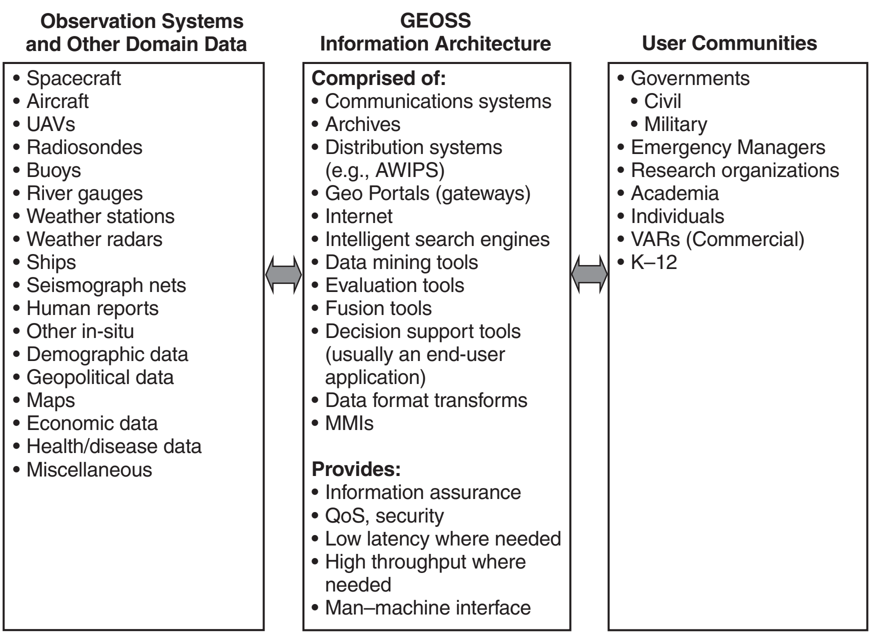


FIGURE 14.26 GEOSS end-to-end architecture.

data from other systems, although in a few cases this is not true, as in some of the National Oceanographic and Atmospheric Administration (NOAA) applications. Another key issue is that many countries simply are not open to sharing their data with the world, even though there are obvious advantages to doing so. They may feel their national security or exclusive economic zones (the 200 nautical mile offshore areas from most countries) are at risk. These issues will need to be resolved before a working implementation can occur. Once these issues are resolved, it is obvious how powerful such a system of systems will be.

A Background on Systems Dynamics

In order to perform a strategic analysis of the GEOSS system, we need to apply Monte Carlo simulation, real options analysis, and couple them with a systems dynamics model. Therefore a quick segue is required here to briefly explain the basics of systems dynamics.

Although systems engineering is a disciplined approach to identifying and specifying requirements as well as architecting systems, systems dynamics allows one to observe the behavior of a system under given circumstances. One such model that makes this possible is the Ventana Vensim model, which allows one to conceptualize, document, simulate, analyze, and optimize models of dynamic systems. Systems dynamics models allow models to be built from causal loops or stocks and flow diagrams. By connecting words with arrows, relationships among system variables are entered and recorded as causal connections. This information is used by the mathematical equations in the model to help form a complete simulation model. The model can be analyzed through the building process, looking at the causes and uses of a variable, and the loops involving the variable. When you have built a model that can be simulated, systems dynamics let you thoroughly explore the behavior of the model.

As a simple example, Figure 14.27 shows the rabbit and fox population behavior and the interaction between the two populations within a systems dynamics model. The model has slider bars built into the birth rates, the initial population, and the average life for the rabbit and fox populations, as well as the fox food requirements and the carrying capacity of the rabbit population. As these bars are adjusted, the remaining variables change the number of births (population and deaths of rabbits and foxes, rabbit crowding, fox consumption of rabbits, and fox food availability). Variables can also be expressed as lookup tables. In this way, we can investigate the behavior of the rabbit and fox population and their interrelationships.

Of course, a model is only as good as its builder and the underlying assumptions. However, systems dynamics have built-in tools that help the builder assess if the model makes sense and the units are correct.

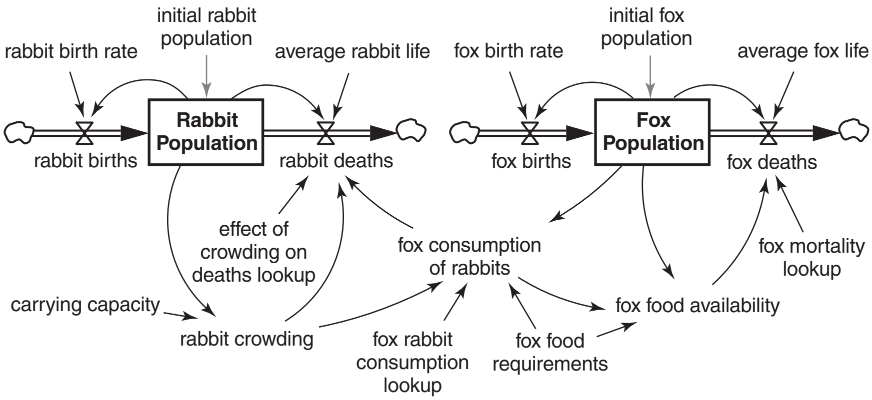


FIGURE 14.27 Sample fox-rabbit population systems dynamics model.

Creation of GEOSS Systems Dynamics Model

Next, the GEOSS model was created using systems dynamics concepts and based on the U.S. military's Office of Force Transformation's model of Network Centric Operations (NCO) as shown in Figure 14.28. The tenets of the NCO model were then modeled. When presented to the GEOSS experts, it was noticed that the tenets of NCO could be slightly modified to fit the GEOSS model, as shown in Figure 14.29.

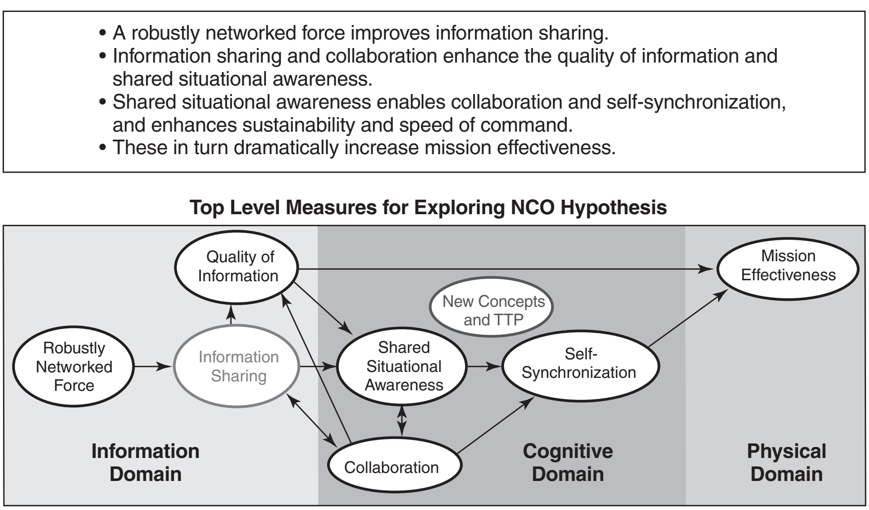


FIGURE 14.28 The Office of Force Transformation NCO model tenets.

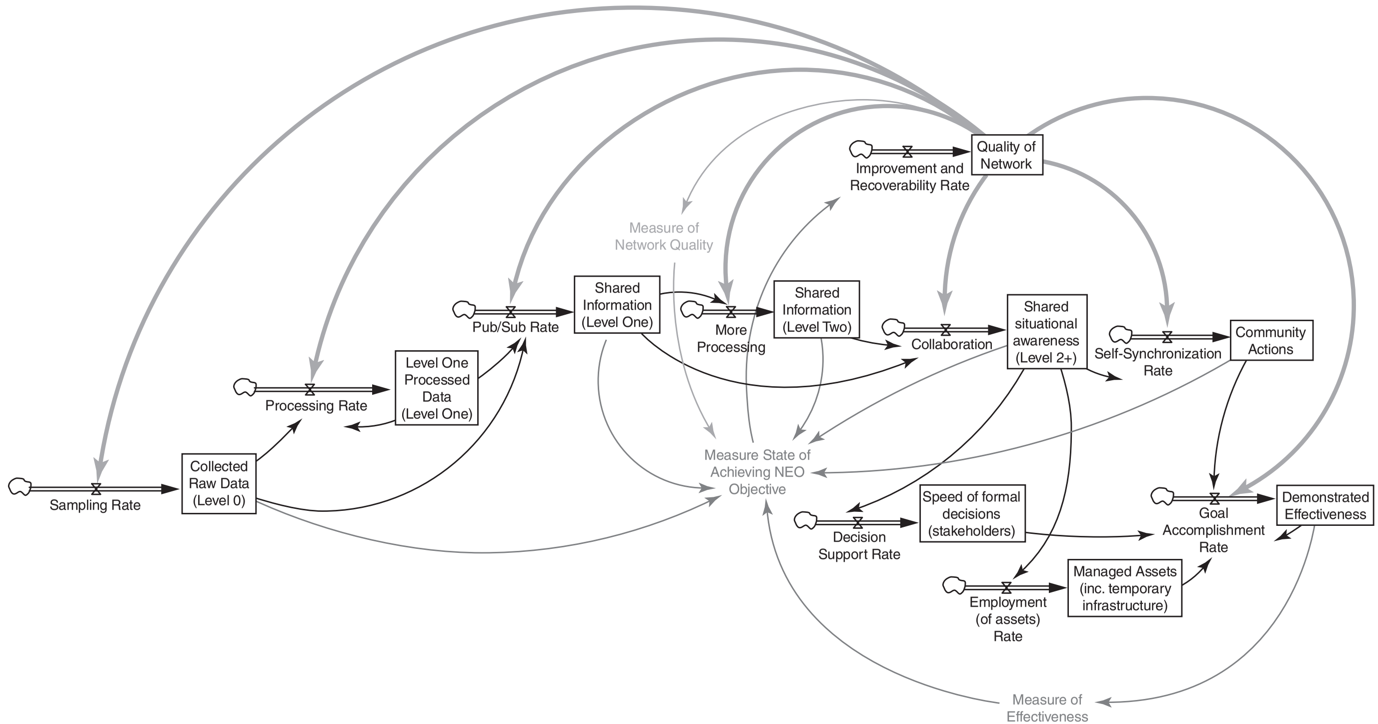


FIGURE 14.29 Systems dynamics model of GEOSS.

Seventy-four technology areas were defined to be required for this large-scale System of Systems (SoS) architecture to operate. An optimization run was then done to determine the most influential technologies in determining system effectiveness, which is largely driven by collaboration. Next, the link to real options analysis was accomplished so the relative value of each technology area could be assessed.

Real Options Valuation Integration and Cost–Benefit Results

The social and economic benefits of a fully developed GEOSS system are very substantial. For instance, according to the U.S. Environmental Protection Agency (EPA), the following is only a small list of the potential benefits:

- We could more accurately know the severity of next winter's weather, with strong implications for emergency managers, transportation, energy and medical personnel, farmers, families, manufacturers, store owners, and others. Weather- and climate-sensitive industries account for one-third of the nation's gross domestic product (GDP), or $3 trillion.
- We could forecast weather one degree Fahrenheit more accurately, saving at least $1 billion annually in U.S. electricity costs.
- With coastal storms reflecting 71 percent, or $7 billion, of U.S. disaster losses every year, improved forecasting would have a major favorable impact on preparedness.
- In the United States, at a cost of $4 billion annually, weather is responsible for about two-thirds of aviation delays—$1.7 billion of which would be avoidable with better observations and forecasts.
- Benefits from more effective air quality monitoring could provide real-time information as well as accurate forecasts that, days in advance, could enable us to mitigate the effects of poor quality through proper transportation and energy use.
- Benefits from ocean instrumentation that, combined with improved satellite earth-observing coverage, could provide revolutionary worldwide and regional climate forecasts, enabling us, for example, to predict years of drought.
- Benefits from real-time monitoring and forecasting of the water quality in every watershed and accompanying coastal areas could provide agricultural interests with immediate feedback and forecasts of the correct amount of fertilizers and pesticides to apply to maximize crop generation at minimum cost, helping to support both healthy ecosystems and greatly increased U.S. fishery output and value from coastal tourism.
- Globally, an estimated 300 million to 500 million people worldwide are infected with malaria each year and about one million die from this

largely preventable disease. With a linked international system, we could pinpoint where the next outbreak of SARS, or bird flu, or West Nile virus, or malaria is likely to hit.

- Natural hazards such as earthquakes, volcanoes, landslides, floods, wildfires, extreme weather, coastal hazards, sea ice and space weather, plus major pollution events, impose a large burden on society. In the United States, the economic cost of disasters averages tens of billions of dollars per year. Disasters are a major cause of loss of life and property. The ability of GEOSS to predict, monitor, and respond to natural and technological hazards is a key consideration in reducing the impact of disasters.

Currently, thousands of individual pieces of technology are gathering earth observations globally. These individual pieces of technologies are demonstrating their value in estimating crop yields, monitoring water and air quality, and improving airline safety. For instance, according to the EPA, U.S. farmers gain about $15 of value for each $1 spent on weather forecasting. Benefits to U.S. agriculture from altering planting decisions are estimated at more than $250 million. The annual economic return to the United States from NOAA's El Niño ocean-observing and forecast system is between 13 and 26 percent. In the meantime, there are thousands of moored and free-floating data buoys in the world's oceans, thousands of land-based environmental stations, and more than 50 environmental satellites orbiting the globe, all providing millions of data sets, but most of these technologies do not yet talk to each other. Until they do, as in a comprehensive GEOSS system, there will always be blind spots and scientific uncertainty. Scientists really cannot know what is happening on our planet without taking the earth's pulse everywhere it beats, all around the globe. Therefore, the challenge is to connect the scientific dots—to build a system of systems that will yield the science on which sound policy must be built.

Strategic Option Pathways

Due to the nature and scope of the project being a global effort, this case study does not expound on all the numerical analyses involved in the quantification of the strategic real options and risk analysis currently being performed. However, a sample strategic tree used for framing real options analysis is provided next to illustrate some of the potential options that GEOSS has. Of course, the entire universe of strategic option pathways and courses of action are a lot more significant than the simple examples illustrated next. See the author's other books on real options for more details on generating strategy trees, as well as modeling and quantifying the real options values using the SLS software (e.g., see *Real Options Analysis: Tools and Techniques, Second Edition* (Wiley, 2005) by Dr. Johnathan Mun).

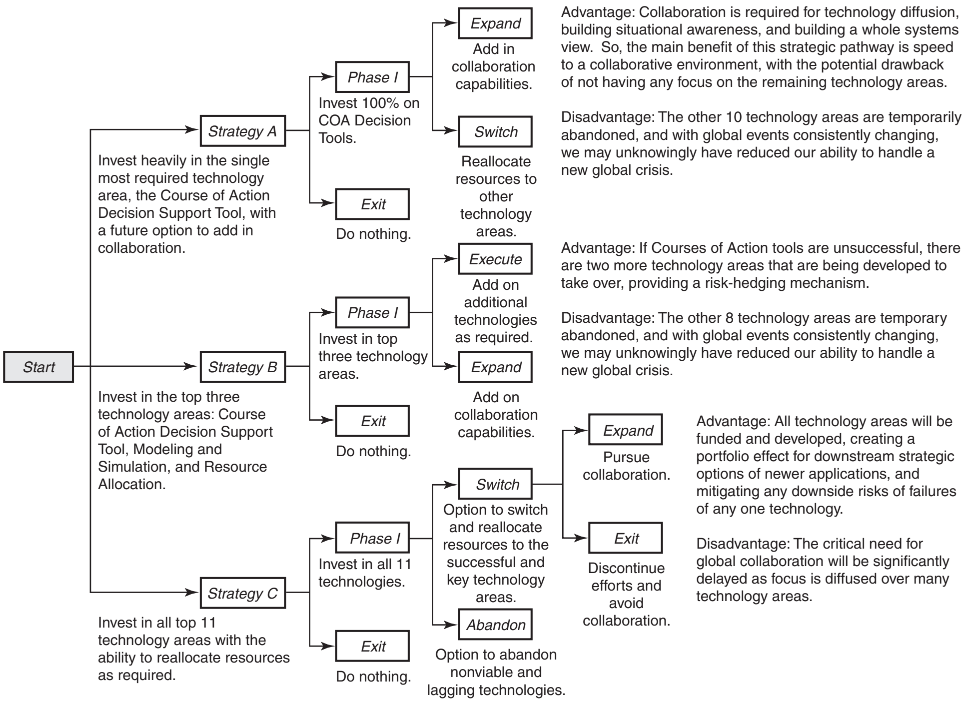


FIGURE 14.30 Sample real options strategies for GEOSS.

To illustrate the basics of the GEOSS options, Figure 14.30 shows three sample pathways of the technology development required as part of a global earth observation system.

Strategy A is to invest heavily in the single most required technology area, the Courses of Action Decision Support tools. It has been determined that Courses of Action technology development builds future options at a faster rate than other technology development because of their influence on collaboration. Collaboration is required for technology diffusion, building situational awareness, and building a whole systems view. The main benefit of this strategic pathway is the speed to a collaborative environment, with the potential drawback of not having any focus on the remaining technology areas.

Strategy B is to invest in the top three technology areas, namely, the Courses of Action Decision Support tools; Modeling and Simulation Decision Support tools; and Resource Allocation tools. One set of technology combinations enables the development of certain follow-on options and activities. So, if Courses of Action tools are unsuccessful, there are two more technology areas that are being developed to take over, providing a risk-hedging mechanism. However, the disadvantage is that the other 8 technology areas are temporarily abandoned, and with global events consistently changing, we may unknowingly have reduced our ability to handle a new global crisis.

Strategy C is to invest in the top 11 technology areas but scale them so that more important technologies get a proportionately higher percentage of the overall investment funding. The advantage is that all technology areas will be funded and developed, creating a portfolio effect for downstream strategic options of newer applications, and mitigating any downside risks of failures of any one technology. However, the disadvantage is that the critical need for global collaboration will be significantly delayed as focus is diffused over many technology areas.

Each of these simple example strategic paths has exit points and each also has an option of whether the technology should be tackled in-house or by some large integrator such as The Boeing Company or by smaller vendors with other expertise in these areas. These are nested options or options within options.

Of course the efforts are ongoing and would pose rather significant analytical and resource challenges. However, with the combinations of simulation, real options, systems dynamics, and optimization tools, the analysis methodology and results can become more valid and robust.

CASE STUDY: VALUING EMPLOYEE STOCK OPTIONS UNDER THE 2004 FAS 123R

This case study is based on Dr. Johnathan Mun's Valuing Employee Stock Options: Under 2004 FAS 123R *(Wiley Finance, 2004). This case study and book applies the same software FASB used to create the valuation examples in FAS 123R's section A87. It was this software application and the training seminars provided by the author for the Board of Directors at FASB, and one-on-one small group trainings for the project managers and research fellows at FASB, that convinced FASB of the pragmatic applications of employee stock options (ESO) valuation. The author consulted for and taught FASB about ESO valuation and is also the creator of the ESO Valuation Toolkit software used by FASB as well as many corporations and consultants.*

Executive Summary

In what the *Wall Street Journal* calls "among the most far-reaching steps that the Financial Accounting Standards Board (FASB) has made in its 30 year history,"[8] in December 2004 FASB released a final revised Statement of Financial Accounting Standard 123 (FAS 123R, or simply denoted as FAS 123) on Share-Based Payment amending the old FAS 123 and 95 issued in October 1995.[9] Basically, the proposal states that starting June 15, 2005, all new and portions of existing employee stock option (ESO) awards that have not yet vested will have to be expensed. In anticipation of the Standard, many companies such as GE and Coca-Cola had already voluntarily expensed their ESOs at the time of writing, while hundreds of other firms were scrambling to look into valuing their ESOs.

The goal of this case study is to provide the reader a better understanding of the valuation applications of FAS 123's preferred methodology—the binomial lattice—through a systematic and objective assessment of the methodology and comparing its results with the Black–Scholes model (BSM). This case study shows that, with care, FAS 123 valuation can be implemented accurately. The analysis performed uses a customized binomial lattice that takes into account real-life conditions such as vesting, employee suboptimal exercise behavior, forfeiture rates, and blackouts, as well as changing dividends, risk-free rates, and volatilities over the life of the ESO. This case study introduces the FAS 123 concept, followed by the different ESO valuation methodologies (closed-form BSM, binomial lattices, and Monte Carlo simulation) and their impacts on valuation. It is shown here that by using the right methodology that still conforms to the FAS 123 requirements, firms can potentially reduce their expenses by millions of dollars a year by avoiding the unnecessary overvaluation of the naïve BSM, using instead a modified and customized binomial lattice model that takes into account suboptimal exercise behavior, forfeiture rates, vesting, blackout dates, and changing inputs over time.

Introduction

The binomial lattice is the preferred method of calculating the fair-market valuation of ESOs in the FAS 123 requirements, but critics argue that companies do not necessarily have the resources in-house or the data availability to perform complex valuations that are both consistent with these new requirements and still be able to pass an audit. Based on a prior published study by the author that was presented to the FASB Board in 2003, it is concluded that the BSM, albeit theoretically correct and elegant, is insufficient and inappropriately applied when it comes to quantifying the fair-market value of ESOs.[10] This is because the BSM is applicable only to European options without dividends, where the holder of the option can exercise the option only on its maturity date and the underlying stock does not pay any dividends.[11] However, in reality, most ESOs are American-type[12] options with dividends, where the option holder can execute the option at any time up to and including the maturity date while the underlying stock pays dividends. In addition, under real-world conditions, ESOs have a time to *vesting* before the employee can execute the option, which may also be contingent on the firm and/or the individual employee attaining a specific performance level (e.g., profitability, growth rate, or stock price hitting a minimum barrier before the options become live), and subject to *forfeitures* when the employee leaves the firm or is terminated prematurely before reaching the vested period. In addition, certain options follow a *tranching* or graduated scale, where a certain percentage of the stock option grants become exercisable every year.[13] Also, employees exhibit erratic exercise behavior where the option will be executed only if it exceeds a particular multiple of the strike price; this is termed the *suboptimal exercise behavior multiple*. Next, the option value may be sensitive to the expected economic environment, as characterized by the term structure of interest rates (i.e., the U.S. Treasuries yield curve) where the risk-free rate changes during the life of the option. Finally, the firm may undergo some corporate restructuring (e.g., divestitures, or mergers and acquisitions that may require a stock swap that changes the volatility of the underlying stock). All these real-life scenarios make the BSM insufficient and inappropriate when used to place a fair-market value on the option grant.[14] In summary, firms can implement a variety of provisions that affect the fair value of the options. The closed-form models such as the BSM or the Generalized Black–Scholes (GBM)—the latter accounts for the inclusion of dividend yields—are inflexible and cannot be modified to accommodate these real-life conditions. Hence, the binomial lattice approach is preferred.

Under very specific conditions (European options without dividends) the binomial lattice and Monte Carlo simulation approaches yield identical values to the BSM, indicating that the two former approaches are robust and exact at the limit. However, when specific real-life business conditions are modeled (i.e., probability of forfeiture, probability the employee leaves or is terminated, time-vesting, suboptimal exercise behavior, and so forth), only the binomial lattice with its highly flexible nature will provide the true fair-market value of the ESO. The BSM takes into account only the following inputs: stock price, strike price, time to maturity, a single risk-free rate, and a single volatility. The GBM accounts for the same inputs as well as a single dividend rate. Hence, in accordance to the FAS 123 requirements, the BSM and GBM fail to account for real-life conditions. In contrast, the binomial lattice can be customized to include the stock price, strike price, time to maturity, a single risk-free rate and/or multiple risk-free rates changing over time, a single volatility and/or multiple volatilities changing over time, a single dividend rate and/or multiple dividend rates changing over time, plus all the other real-life factors including, but not limited to, vesting periods, suboptimal early exercise behavior, blackout periods, forfeiture rates, stock price and performance barriers, and other exotic contingencies. Note that the binomial lattice results revert to the GBM if these real-life conditions are negligible.

The two most important and convincing arguments for using binomial lattices are (1) that FASB requires it and states that the binomial lattice is the preferred method for ESO valuation and (2) that lattices can substantially reduce the cost of the ESO by more appropriately mirroring real-life conditions. Here is a sample of FAS 123's requirements discussing the use of binomial lattices.

> *B64. As discussed in paragraphs A10–A17, closed-form models are one acceptable technique for estimating the fair value of employee share options. However, a lattice model (or other valuation technique, such as a Monte Carlo simulation technique, that is not based on a closed-form equation) can accommodate the term structures of risk-free interest rates and expected volatility, as well as expected changes in dividends over an option's contractual term. A lattice model also can accommodate estimates of employees' option exercise patterns and post-vesting employment termination during the option's contractual term, and thereby can more fully reflect the effect of those factors than can an estimate developed using a closed-form model and a single weighted-average expected life of the options.*
>
> *A15. The Black–Scholes–Merton formula assumes that option exercises occur at the end of an option's contractual term, and that expected volatility, expected dividends, and risk-free interest rates are constant over the option's term. If used to estimate the fair value of instruments in the scope of this Statement, the Black–Scholes–Merton formula must be adjusted to take account of certain characteristics of employee share options and similar instruments that are not consistent with the model's assumptions (for example, the ability to exercise before the end of the option's contractual term). Because of the nature of the formula, those adjustments take the form of weighted average assumptions about those characteristics. In contrast, a lattice model can be designed to accommodate dynamic assumptions of expected volatility and dividends over the option's contractual term, and estimates of expected option exercise patterns during the option's contractual term, including the effect of blackout periods. Therefore, the design of a lattice model more fully reflects the substantive characteristics of a particular employee share option or similar instrument. Nevertheless, both a lattice model and the Black–Scholes–Merton formula, as well as other valuation techniques that meet the requirements in paragraph A8, can provide a fair value estimate that is consistent with the measurement objective and fair-value-based method of this Statement. However, if an entity uses a lattice*

model that has been modified to take into account an option's contractual term and employees' expected exercise and post-vesting employment termination behavior, the expected term is estimated based on the resulting output of the lattice. For example, an entity's experience might indicate that option holders tend to exercise their options when the share price reaches 200 percent of the exercise price. If so, that entity might use a lattice model that assumes exercise of the option at each node along each share price path in a lattice at which the early exercise expectation is met, provided that the option is vested and exercisable at that point. Moreover, such a model would assume exercise at the end of the contractual term on price paths along which the exercise expectation is not met but the options are in-the-money at the end of the contractual term. That method recognizes that employees' exercise behavior is correlated with the price of the underlying share. Employees' expected post-vesting employment termination behavior also would be factored in. Expected term, which is a required disclosure (paragraph A240), then could be estimated based on the output of the resulting lattice.

In fact, some parts of the FAS 123 Final Requirements cannot be modeled with a traditional Black–Scholes model. A lattice is required to model items such as suboptimal exercise behavior multiple, forfeiture rates, vesting, blackout periods, and so forth. This case study and the software used to compute the results use both a binomial (and trinomial) lattice as well as closed-form Black–Scholes models to compare the results. The specific FAS 123 paragraphs describing the use of lattices include:

A27. However, if an entity uses a lattice model that has been modified to take into account an option's contractual term and employees' expected exercise and post-vesting employment termination behavior, the expected term is estimated based on the resulting output of the lattice. For example, an entity's experience might indicate that option holders tend to exercise their options when the share price reaches 200 percent of the exercise price. If so, that entity might use a lattice model that assumes exercise of the option at each node along each share price path in a lattice at which the early exercise expectation is met, provided that the option is vested and exercisable at that point.

A28. Other factors that may affect expectations about employees' exercise and post-vesting employment termination behavior include the following:

a. The vesting period of the award. An option's expected term must at least include the vesting period.
b. Employees' historical exercise and post-vesting employment termination behavior for similar grants.

c. Expected volatility of the price of the underlying share.
d. Blackout periods and other coexisting arrangements such as agreements that allow for exercise to automatically occur during blackout periods if certain conditions are satisfied.
e. Employees' ages, lengths of service, and home jurisdictions (that is, domestic or foreign).

Therefore, based on the preceding justifications, and in accordance to the requirements and recommendations set forth by the revised FAS 123, which prefers the binomial lattice, it is hereby concluded that the customized binomial lattice is the best and preferred methodology to calculate the fair-market value of ESOs.

Application of the Preferred Method

In applying the customized binomial lattice methodology, several inputs have to be determined:

- Stock price at grant date.
- Strike price of the option grant.
- Time to maturity of the option.
- Risk-free rate over the life of the option.
- Dividend yield of the option's underlying stock over the life of the option.
- Volatility over the life of the option.
- Vesting period of the option grant.
- Suboptimal exercise behavior multiples over the life of the option.
- Forfeiture and employee turnover rates over the life of the option.
- Blackout dates postvesting when the options cannot be exercised.

The analysis assumes that the employee cannot exercise the option when it is still in the vesting period. Further, if the employee is terminated or decides to leave voluntarily during this vesting period, the option grant will be forfeited and presumed worthless. In contrast, after the options have been vested, employees tend to exhibit erratic exercise behavior where an option will be exercised only if it breaches the suboptimal exercise behavior multiple.[15] However, the options that have vested must be exercised within a short period if the employee leaves voluntarily or is terminated, regardless of the suboptimal behavior threshold—that is, if forfeiture occurs (measured by the historical option forfeiture rates as well as employee turnover rates). Finally, if the option expiration date has been reached, the option will be exercised if it is in-the-money, and expire worthless if it is at-the-money or out-of-the-money. The next section details the results obtained from such an analysis.

ESO Valuation Toolkit Software

It is theoretically impossible to solve a large binomial lattice ESO valuation without the use of software algorithms.[16] The analyses results in this case study were performed using the author's Employee Stock Options Valuation Toolkit 1.1 software (Figure 14.31), which is the same software used by FASB to convince itself that ESO valuation is pragmatic and manageable. In fact, FASB used this software to calculate the valuation example in the Final FAS 123 release in sections A87–A88 (illustrated later). Figure 14.32 shows a sample module for computing the Customized American Option using binomial lattices with vesting, forfeiture rate, suboptimal exercise behavior multiple, and changing risk-free rates and volatilities over time. The Real Options Super Lattice Solver software also can be used to create any customized ESO model using binomial lattices, FASB's favored method.

The software shows the applications of both closed-form models such as the BSM/GBM and binomial lattice methodologies. By using binomial lattice methodologies, more complex ESOs can be solved. For instance, the Customized Advanced Option (Figure 14.32) shows how multiple variables can be varied over time (risk-free, dividend, volatility, forfeiture rate, suboptimal exercise behavior multiple, and so forth). In addition, for added flexibility, the Super Lattice Solver module allows the expert user to create and solve his

FIGURE 14.31 ESO Valuation Toolkit 1.1 software.

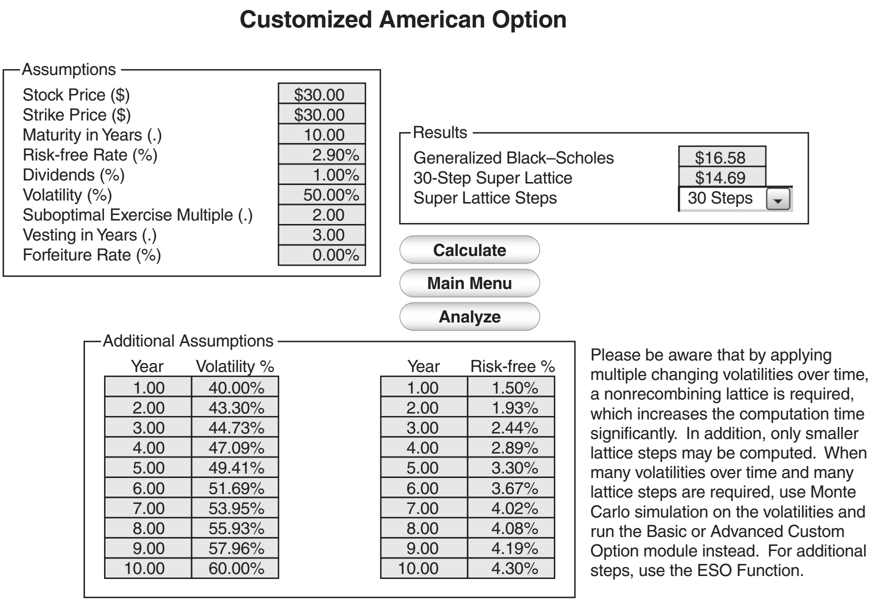


FIGURE 14.32 Customized advanced option model.

or her own customized ESO. This feature allows management to experiment with different flavors of ESO as well as to engineer one that would suit its needs, by balancing fair and equitable value to employees, with cost minimization to its shareholders.

Figure 14.32 shows the solution of the case example provided in section A87 of the Final 2004 FAS 123 standards. Specifically, A87–A88 state:

> *A87. The following table shows assumptions and information about the share options granted on January 1, 20X5.*
>
> *Share options granted 900,000*
> *Employees granted options 3,000*
> *Expected forfeitures per year 3.0%*
> *Share price at the grant date $30*
> *Exercise price $30*
> *Contractual term (CT) of options 10 years*
> *Risk-free interest rate over CT 1.5% to 4.3%*
> *Expected volatility over CT 40% to 60%*
> *Expected dividend yield over CT 1.0%*
> *Suboptimal exercise factor 2*

A88. This example assumes that each employee receives an equal grant of 300 options. Using as inputs the last 7 items from the table above, Entity T's lattice-based valuation model produces a fair value of \$14.69 per option. A lattice model uses a suboptimal exercise factor to calculate the expected term (that is, the expected term is an output) rather than the expected term being a separate input. If an entity uses a Black–Scholes–Merton option-pricing formula, the expected term would be used as an input instead of a suboptimal exercise factor.

Figure 14.32 shows the result as \$14.69, the answer that FASB uses in its example. The forfeiture rate of 3 percent used by FASB's example is applied outside of the model to discount for the quantity reduced over time. The software allows the ability to input the forfeiture rates (pre- and post-vesting) inside or outside of the model. In this specific example, we set forfeiture rate to zero in Figure 14.32 and adjust the quantity outside, just as FASB does, in A91:

The number of share options expected to vest is estimated at the grant date to be 821,406 ($900,000 \times .97^3$).

In fact, using the ESO Valuation Toolkit software and Excel's goal seek function, we can find that the expected life of this option is 6.99 years. We can then justify the use of 6.99 years as the input into a modified GBM to obtain the same result at \$14.69, something that cannot be done without the use of the binomial lattice approach.

Technical Justification of Methodology Employed

This section illustrates some of the technical justifications that make up the price differential between the GBM and the customized binomial lattice models. Figure 14.33 shows a tornado chart and how each input variable in a customized binomial lattice drives the value of the option.[17] Based on the chart, it is clear that volatility is not the single key variable that drives option value. In fact, when vesting, forfeiture, and suboptimal behavior elements are added to the model, their effects dominate that of volatility. The chart illustrated is based on a typical case and cannot be generalized across all cases.

In contrast, volatility is a significant variable in a simple BSM as can be seen in Figure 14.34. This is because there is less interaction among input variables due to the fewer input variables, and for most ESOs that are issued at-the-money, volatility plays an important part when there are no other dominant inputs.

In addition, the interactions among these new input variables are nonlinear. Figure 14.35 shows a spider chart[18] where it can be seen that vesting, forfeiture rates, and suboptimal exercise behavior multiples have nonlinear

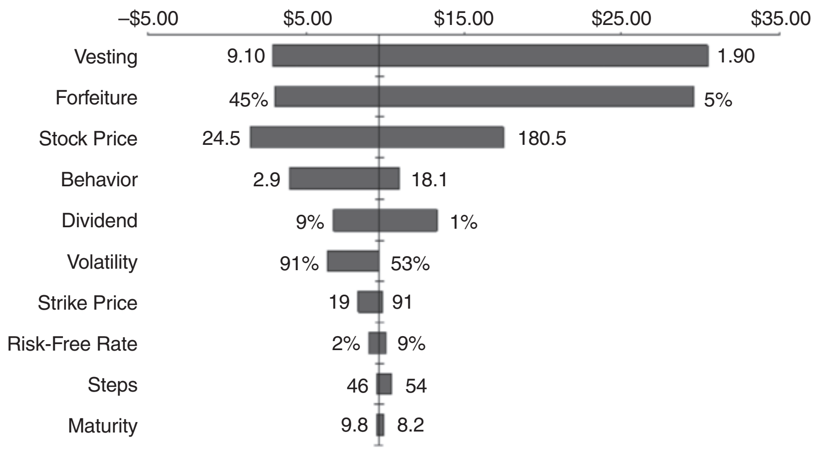


FIGURE 14.33 Tornado chart listing the critical input factors of a customized binomial model.

effects on option value. That is, the lines in the spider chart are not straight but curve at certain areas, indicating that there are nonlinear effects in the model. This means that we cannot generalize these three variables' effects on option value (for instance, we cannot generalize that if a 1 percent increase in forfeiture rate will decrease option value by 2.35 percent, it means that a 2 percent increase in forfeiture rate drives option value down 4.70 percent, and so forth). This is because the variables interact differently at different

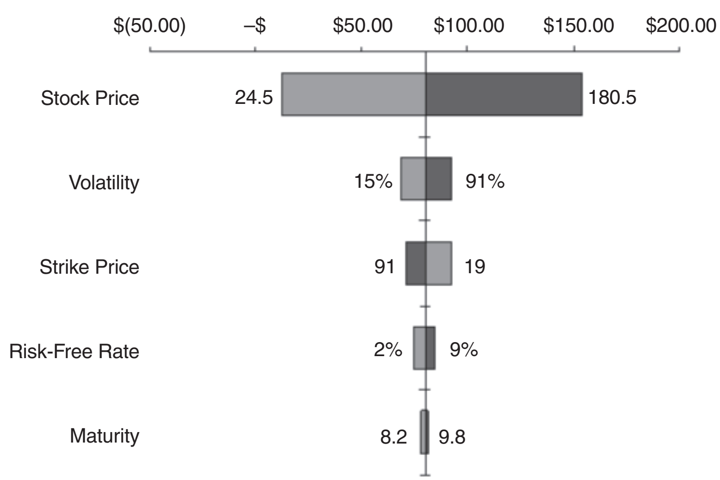


FIGURE 14.34 Tornado chart listing the critical input factors of the BSM.

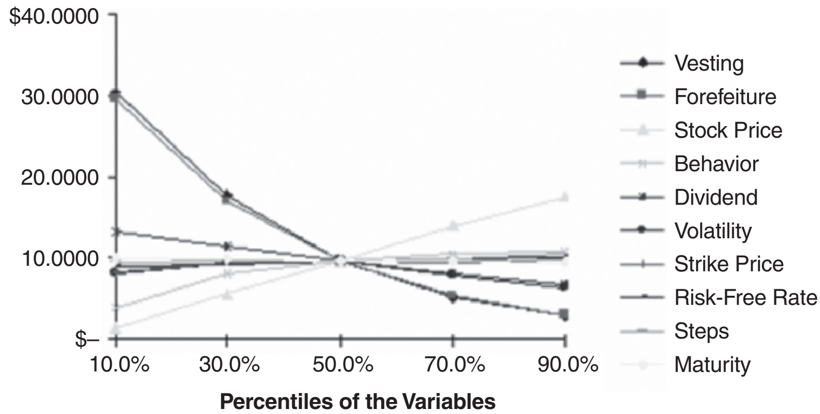


FIGURE 14.35 Spider chart showing the nonlinear effects of input factors in the binomial model.

input levels. The conclusion is that we really cannot say a priori what the direct effects are of changing one variable on the magnitude of the final option value. More detailed analysis will have to be performed in each case.

Although the tornado and spider charts illustrate the impact of each input variable on the final option value, the effects are static; that is, one variable is tweaked at a time to determine its ramifications on the option value. However, as shown, the effects are sometimes nonlinear, which means we need to change all variables simultaneously to account for their interactions. Figure 14.36 shows a Monte Carlo simulated dynamic sensitivity

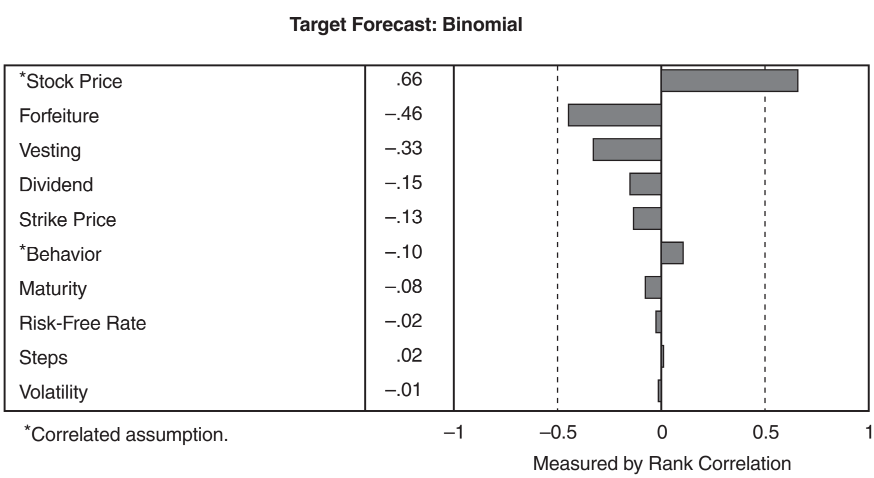


FIGURE 14.36 Dynamic sensitivity with simultaneously changing input factors in the binomial model.

chart where forfeiture, vesting, and suboptimal exercise behavior multiple are determined to be important variables, while volatility is again relegated to a less important role. The dynamic sensitivity chart perturbs all input variables simultaneously for thousands of trials, and captures the effects on the option value. This approach is valuable in capturing the net interaction effects among variables at different input levels.

From this preliminary sensitivity analysis, we conclude that incorporating forfeiture rates, vesting, and suboptimal exercise behavior multiple is vital to obtaining a fair-market valuation of ESOs due to their significant contributions to option value. In addition, we cannot generalize each input's effects on the final option value. Detailed analysis has to be performed to obtain the option's value every time.

Options with Vesting and Suboptimal Behavior

Further investigation into the elements of suboptimal behavior[19] and vesting yields the chart shown in Figure 14.37. Here we see that at lower suboptimal exercise behavior multiples (within the range of 1 to 6), the stock option value can be significantly lower than that predicted by the BSM. With a 10-year vesting stock option, the results are identical regardless of the suboptimal exercise behavior multiple—its flat line bears the same value as the BSM result. This is because for a 10-year vesting of a 10-year maturity option, the option reverts to a perfect European option, where it can be exercised only at expiration. The BSM provides the correct result in this case.

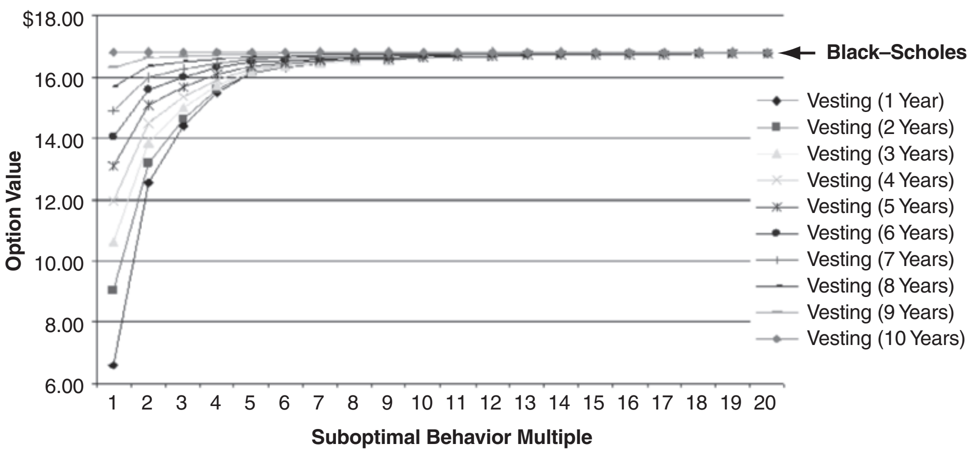


FIGURE 14.37 Impact of suboptimal exercise behavior and vesting on option value in the binomial model. (Assumptions used: stock and strike price of $25, 10-year maturity, 5% risk-free rate, 50% volatility, 0% dividends, suboptimal exercise behavior multiple range of 1–20, vesting period of 1–10 years, and tested with 100–5,000 binomial lattice steps.)

However, when suboptimal exercise behavior multiple is low, the option value decreases because employees holding the option will tend to exercise the option suboptimally—that is, the option will be exercised earlier and at a lower stock price than optimal. Hence, the option's upside value is not maximized. As an example, suppose an option's strike price is $10 while the underlying stock is highly volatile. If an employee exercises the option at $11 (this means a 1.10 suboptimal exercise multiple), he or she may not be capturing the entire upside potential of the option as the stock price can go up significantly higher than $11 depending on the underlying volatility. Compare this to another employee who exercises the option when the stock price is $20 (suboptimal exercise multiple of 2.0) versus one who does so at a much higher stock price. Thus, lower suboptimal exercise behavior means a lower fair-market value of the stock option. This suboptimal exercise behavior has a higher impact when stock prices at grant date are forecast to be high. Figure 14.38 shows that (at the lower end of the suboptimal multiples) a steeper slope occurs the higher the initial stock price at grant date.

Figure 14.39 shows that for higher volatility stocks, the suboptimal region is larger and the impact to option value is greater, but the effect is gradual. For instance, for the 100 percent volatility stock, the suboptimal region extends from a suboptimal exercise behavior multiple of 1.0 to approximately 9.0 versus from 1.0 to 2.0 for the 10 percent volatility stock. In addition, the vertical distance of the 100 percent volatility stock extends from

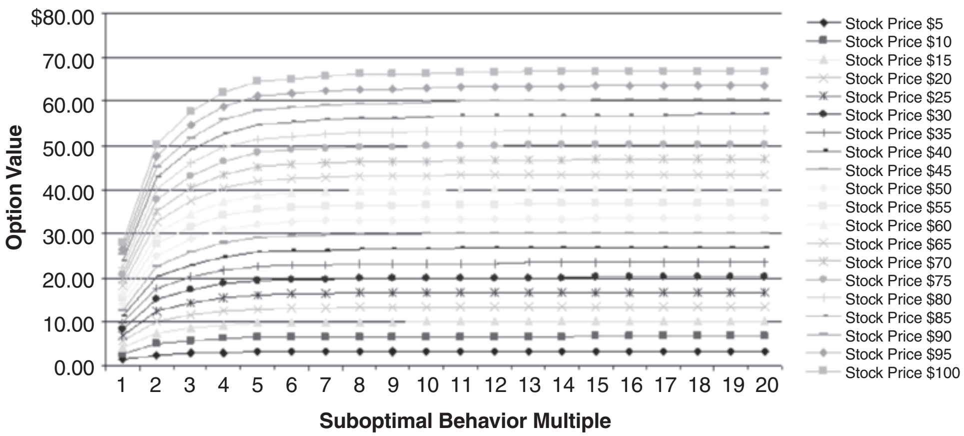


FIGURE 14.38 Impact of suboptimal exercise behavior and stock price on option value in the binomial model. (Assumptions used: stock and strike price range of $5 to $100, 10-year maturity, 5% risk-free rate, 50% volatility, 0% dividends, suboptimal exercise behavior multiple range of 1–20, 4-year vesting, and tested with 100–5,000 binomial lattice steps.)

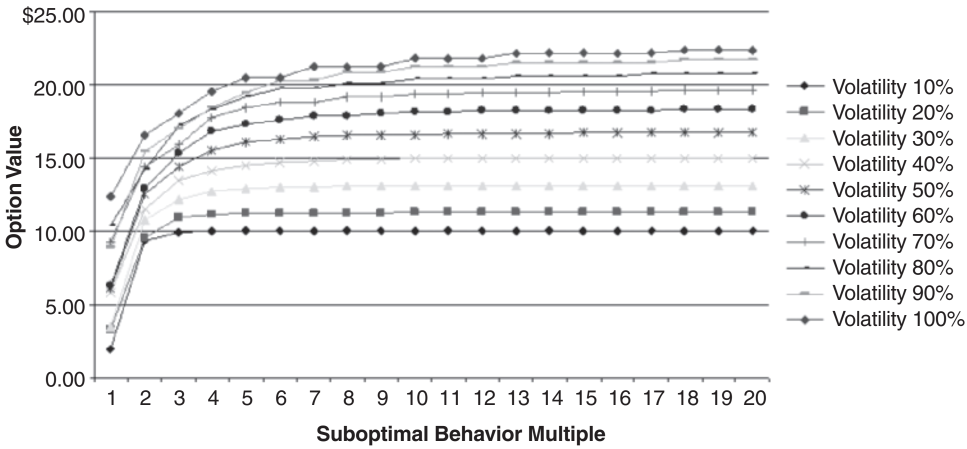


FIGURE 14.39 Impact of suboptimal exercise behavior and volatility on option value in the binomial model. (Assumptions used: stock and strike price of $25, 10-year maturity, 5% risk-free rate, 10–100% volatility range, 0% dividends, suboptimal exercise behavior multiple range of 1–20, 1-year vesting, and tested with 100–5,000 binomial lattice steps.)

$12 to $22 with a $10 range, as compared to $2 to $10 with an $8 range for the 10 percent volatility stock. Therefore, the higher the stock price at grant date and the higher the volatility, the greater the impact of suboptimal behavior will be on the option value. *In all cases,* the BSM results are the horizontal lines in the charts (Figures 14.38 and 14.39). That is, the BSM will always generate the maximum option value assuming optimal behavior, and overexpense the option significantly. A GBM or BSM cannot be modified to account for this suboptimal exercise behavior; only the binomial lattice can be used.

Options with Forfeiture Rates

Figure 14.40 illustrates the reduction in option value when the forfeiture rate increases. The rate of reduction changes depending on the vesting period. The longer the vesting period, the more significant the impact of forfeitures will be, illustrating once again the nonlinear interacting relationship between vesting and forfeitures (i.e., the lines in Figure 14.40 are curved and nonlinear). This is intuitive because the longer the vesting period, the lower the compounded probability that an employee will still be employed in the firm and the higher the chances of forfeiture, reducing the expected value of the option.

Again, we see that the BSM result is the highest possible value assuming a 10-year vesting in a 10-year maturity option with zero forfeiture (Figure

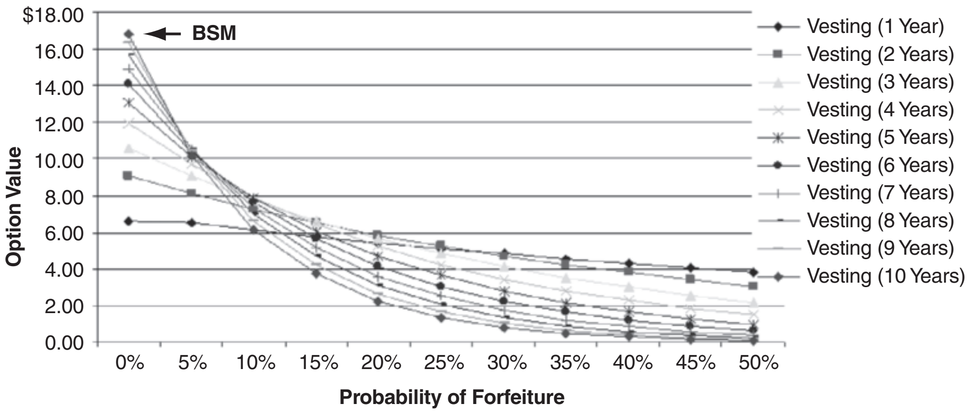


FIGURE 14.40 Impact of forfeiture rates and vesting on option value in the binomial model. (Assumptions used: stock and strike price of $25, 10-year maturity, 5% risk-free rate, 50% volatility, 0% dividends, suboptimal behavior 1.01, vesting period of 1–10 years, forfeiture range 0–50%, and tested with 100–5,000 binomial lattice steps.)

14.40). In addition, forfeiture rates can be negatively correlated to stock price—if the firm is doing well, its stock price usually increases, making the option more valuable and making the employees less likely to leave and the firm less likely to lay off its employees. Because the rate of forfeitures is uncertain (forfeiture rate fluctuations typically occur in the past due to business and economic environments, and will most certainly fluctuate again in the future) and is negatively correlated to the stock price, we can also apply a correlated Monte Carlo simulation on forfeiture rates in conjunction with the customized binomial lattices (shown later in this case study). The BSM will always generate the maximum option value assuming all options will fully vest and will overexpense the option significantly. The ESO Valuation software can account for forfeiture rates, while the accompanying Super Lattice Solver can account for different prevesting and postvesting forfeiture rates in the lattices.

Options Where Risk-Free Rate Changes Over Time

Another input assumption is the risk-free rate. Figure 14.41 illustrates the effects of changing risk-free rates over time on option valuation. When other exotic inputs are added, the changing risk-free lattice model has an overall lower valuation. In addition, due to the time value of money, discounting more heavily in the future will reduce the option's value. In other words, Figure 14.41 compares an upward sloping yield curve, a downward sloping

Basic Input Parameters		Year	Static Base Case	Increasing Risk-Free Rates	Decreasing Risk-Free Rates	Risk-Free Rate Smile	Risk-Free Rate Frown
Stock Price	$100.00	1	5.50%	1.00%	10.00%	8.00%	3.50%
Strike Price	$100.00	2	5.50	3.00	9.00	7.00	4.00
Maturity	10.00	3	5.50	3.00	8.00	5.00	5.00
Volatility	45.00	4	5.50	4.00	7.00	4.00	7.00
Dividend Rate	4.00	5	5.50	5.00	6.00	3.50	8.00
Lattice Steps	1000	6	5.50	6.00	5.00	3.50	8.00
Suboptimal Behavior	1.80	7	5.50	7.00	4.00	4.00	7.00
Vesting Period	4.00	8	5.50	8.00	3.00	5.00	5.00
Forfeiture Rate	10.00	9	5.50	9.00	2.00	7.00	4.00
		10	5.50	10.00	1.00	8.00	3.50
		Average	*5.50*	*5.50*	*5.50*	*5.50*	*5.50*
		BSM using 5.50% Average Rate	$37.45	$37.45	$37.45	$37.45	$37.45
		Forfeiture Modified BSM using 5.50% Average Rate	$33.71	$33.71	$33.71	$33.71	$33.71
		Changing Risk-free Binomial Lattice	$25.92	$24.31	$27.59	$26.04	$25.76

FIGURE 14.41 Effects of changing risk-free rates on option value. These results only illustrate a typical case and should not be generalized across all possible cases.

yield curve, risk-free rate smile, and risk-free rate frown. When the term structure of interest rates increases over time, the option value calculated using a customized changing risk-free rate binomial lattice is lower ($24.31) than that calculated using an average of the changing risk-free rates ($25.92) base case. The reverse is true for a downward-sloping yield curve. In addition, Figure 14.41 shows a risk-free yield curve frown (low rates followed by high rates followed by low rates) and a risk-free yield curve smile (high rates followed by low rates followed by high rates). The results indicate that using a single average rate will overestimate an upward-sloping yield curve, underestimate a downward-sloping yield curve, underestimate a yield curve smile, and overestimate a yield curve frown. Therefore, whenever appropriate, use all available information in terms of forward risk-free rates, one rate for each year.

Options Where Volatility Changes Over Time

Figure 14.42 illustrates the effects of changing volatilities on an ESO. If volatility changes over time, the BSM ($71.48) using the average volatility over time will *always* overestimate the true option value when there are other exotic inputs. In addition, compared to the $38.93 base case, slowly increasing volatilities over time from a low level has lower option values, while a decreasing volatility from high values and volatility smiles and frowns have higher values than using the average volatility estimate.

Basic Input Parameters		Year	Static Base Case	Increasing Volatilities	Decreasing Volatilities	Volatility Smile	Volatility Frown
Stock Price	$100.00	1	55.00%	10.00%	100.00$	80.00%	35.00%
Strike Price	$100.00	2	55.00	20.00	90.00	70.00	40.00
Maturity	10.00	3	55.00	30.00	80.00	50.00	50.00
Risk-free Rate	5.50	4	55.00	40.00	70.00	40.00	70.00
Dividend Rate	0.00	5	55.00	50.00	60.00	35.00	80.00
Lattice Steps	10	6	55.00	60.00	50.00	35.00	80.00
Suboptimal Behavior	1.80	7	55.00	70.00	40.00	40.00	70.00
Vesting Period	4.00	8	55.00	80.00	30.00	50.00	50.00
Forfeiture Rate	10.00	9	55.00	90.00	20.00	70.00	40.00
		10	55.00	100.00	10.00	80.00	35.00
		Average	*55.00*	*55.00*	*55.00*	*55.00*	*55.00*
		BSM using 5.50% Average Rate	$71.48	$71.48	$71.48	$71.48	$71.48
		Forfeiture Modified BSM using 5.50% Average Rate	$64.34	$64.34	$64.34	$64.34	$64.34
		Changing Risk-free Binomial Lattice	$38.93	$32.35	$45.96	$39.56	$39.71

FIGURE 14.42 Effects of changing volatilities on option value.

Options Where Dividend Yield Changes Over Time

Dividend yield is a simple input that can be obtained from corporate dividend policies or publicly available historical market data. It is the total dividend payments computed as a percentage of stock price that is paid out over the course of a year. The typical dividend yield is between 0 percent and 7 percent. In fact, about 45 percent of all publicly traded firms in the United States pay dividends. Of those that pay a dividend, 85 percent have a yield of 7 percent or below, and 95 percent have a yield of 10 percent or below.[20] Dividend yield is an interesting variable with very little interaction with other exotic input variables. It has a close to linear effect on option value, whereas the other exotic input variables do not. For instance, Figure 14.43 illustrates the effects of different maturities on the same option. The higher the maturity, the higher the option value, but the option value increases at a decreasing rate.

In contrast, Figure 14.44 illustrates the near-linear effects of dividends even when some of the exotic inputs have been changed. Whatever the change in variable is, the effects of dividends are always very close to linear. While Figure 14.44 illustrates many options with unique dividend rates, Figure 14.45 illustrates the effects of changing dividends over time on a single option. That is, the results shown in Figure 14.44 are based on comparing different options with different dividend rates, whereas the results shown in Figure 14.45 are based on a single option whose underlying stock's dividend yields are changing over the life of the option.

Maturity	Option Value	Change
1	$25.16	—
2	32.41	28.84%
3	35.35	9.08
4	36.80	4.08
5	37.87	2.91
6	38.41	1.44
7	38.58	0.43

FIGURE 14.43 Nonlinear effects of maturity. (Assumptions used: stock price and strike price are set at $100, 5% risk-free rate, 75% volatility, and 1,000 steps in the customized lattice, 1.8 behavior multiple, 1-year vesting, 10% forfeiture rate.)

	1.8 Behavior Multiple, 4-Year Vesting, 10% Forfeiture Rate		1.8 Behavior Multiple, 1-Year Vesting, 10% Forfeiture Rate		3.0 Behavior Multiple, 1-Year Vesting, 10% Forfeiture Rate	
Dividend Rate	Option Value	Change	Option Value	Change	Option Value	Change
0%	$42.15		$42.41		$49.07	
1	39.94	–5.24%	41.47	–2.20%	47.67	–2.86%
2	37.84	–5.27	40.55	–2.22	46.29	–2.89
3	35.83	–5.30	39.65	–2.24	44.94	–2.92
4	33.92	–5.33	38.75	–2.26	43.61	–2.95
5	32.10	–5.37	37.87	–2.28	42.31	–2.98

	$50 Stock Price, 1.8 Behavior Multiple, 1-Year Vesting, 10% Forfeiture Rate		1.8 Behavior Multiple, 1-Year Vesting, 5% Forfeiture Rate	
Dividend Rate	Option Value	Change	Option Value	Change
0%	$21.20		$45.46	
1	20.74	–2.20%	44.46	–2.20%
2	20.28	–2.22	43.47	–2.23
3	19.82	–2.24	42.49	–2.25
4	19.37	–2.26	41.53	–2.27
5	18.93	–2.28	40.58	–2.29

FIGURE 14.44 Near-linear effects of dividends.

Scenario	Option Value	Change	Notes
Static 3% Dividend	$39.65	0.00%	Dividends are kept steady at 3%
Increasing Gradually	$40.94	3.26%	1% to 5% with 1% increments (average of 3%)
Decreasing Gradually	$38.39	–3.17%	5% to 1% with –1% increments (average of 3%)
Increasing Jumps	$41.70	5.19%	0%, 0%, 5%, 5%, 5% (average of 3%)
Decreasing Jumps	$38.16	–3.74%	5%, 5%, 5%, 0%, 0% (average of 3%)

FIGURE 14.45 Effects of changing dividends over time. (Assumptions used: stock price and strike price are set at $100, 5-year maturity, 5% risk-free rate, 75% volatility, 1,000 steps in the customized lattice, 1.8 behavior multiple, 10% forfeiture rate, and 1-year vesting.)

Clearly, a changing-dividend option has some value to add in terms of the overall option valuation results. Therefore, if the firm's stock pays a dividend, then the analysis should also consider the possibility of dividend yields changing over the life of the option.

Options Where Blackout Periods Exist

Another item of interest is blackout periods, the dates that ESOs cannot be executed. These dates are usually several weeks before and several weeks after an earnings announcement (usually on a quarterly basis). In addition, only senior executives with fiduciary responsibilities have these blackout dates, and, hence, their proportion is relatively small compared to the rest of the firm. Figure 14.46 illustrates the calculations of a typical ESO with different blackout dates. In the case where there are only a few blackout days a month, there is little difference between options with blackout dates and those without blackout dates. In fact, if the suboptimal exercise behavior multiple is small (a 1.8 ratio is assumed in this case), blackout dates

Blackout Dates	*Option Value*
No Blackouts	$43.16
Every 2 years evenly spaced	43.16
First 5 years annual blackouts only	43.26
Last 5 years annual blackouts only	43.16
Every 3 months for 10 years	43.26

FIGURE 14.46 Effects of blackout periods on option value. (Assumptions used: stock and strike price of $100, 75% volatility, 5% risk-free rate, 10-year maturity, no dividends, 1-year vesting, 10% forfeiture rate, and 1,000 lattice steps.)

at strategic times will actually prevent the option holder from exercising suboptimally and sometimes even increase the value of the option ever so slightly.

The analysis shown as Figure 14.46 assumes only a small percentage of blackout dates in a year (e.g., during several days in a year, the ESO cannot be executed). This may be the case for certain so-called brick-and-mortar companies, and, as such, blackout dates can be ignored. However, in other firms such as those in the biotechnology and high-tech industries, blackout periods play a more significant role. For instance, in a biotech firm, blackout periods may extend 4–6 weeks every quarter, straddling the release of its quarterly earnings. In addition, blackout periods prior to the release of a new product may exist. Therefore, the proportion of blackout dates with respect to the life of the option may reach upward of 35–65 percent per year. In such cases, blackout periods will significantly affect the value of the option. For instance, Figure 14.47 illustrates the differences between a customized binomial lattice with and without blackout periods. By adding in the real-life elements of blackout periods, the ESO value is further reduced by anywhere between 10 percent and 35 percent depending on the rate of forfeiture and volatility. As expected, the reduction in value is nonlinear, as the effects of blackout periods will vary depending on the other input variables involved in the analysis.

% Difference between no blackout periods versus significant blackouts	Volatility (25%)	Volatility (30%)	Volatility (35%)	Volatility (40%)	Volatility (45%)	Volatility (50%)
Forfeiture Rate (5%)	–17.33%	–13.18%	–10.26%	–9.21%	–7.11%	–5.95%
Forfeiture Rate (6%)	–19.85%	–15.17%	–11.80%	–10.53%	–8.20%	–6.84%
Forfeiture Rate (7%)	–22.20%	–17.06%	–13.29%	–11.80%	–9.25%	–7.70%
Forfeiture Rate (8%)	–24.40%	–18.84%	–14.71%	–13.03%	–10.27%	–8.55%
Forfeiture Rate (9%)	–26.44%	–20.54%	–16.07%	–14.21%	–11.26%	–9.37%
Forfeiture Rate (10%)	–28.34%	–22.15%	–17.38%	–15.35%	–12.22%	–10.17%
Forfeiture Rate (11%)	–30.12%	–23.67%	–18.64%	–16.45%	–13.15%	–10.94%
Forfeiture Rate (12%)	–31.78%	–25.11%	–19.84%	–17.51%	–14.05%	–11.70%
Forfeiture Rate (13%)	–33.32%	–26.48%	–21.00%	–18.53%	–14.93%	–12.44%
Forfeiture Rate (14%)	–34.77%	–27.78%	–22.11%	–19.51%	–15.78%	–13.15%
Forfeiture Rate (14%)	–34.77%	–27.78%	–22.11%	–19.51%	–15.78%	–13.15%

FIGURE 14.47 Effects of significant blackouts (different forfeiture rates and volatilities). (Assumptions used: stock and strike price range of \$30 to \$100, 45% volatility, 5% risk-free rate, 10-year maturity, dividend range 0–10%, vesting of 1–4 years, 5–14% forfeiture rate, suboptimal exercise behavior multiple range of 1.8–3.0, and 1,000 lattice steps.)

% Difference between no blackout periods versus significant blackouts	Vesting (1)	Vesting (2)	Vesting (3)	Vesting (4)
Dividends (0%)	–8.62%	–6.93%	–5.59%	–4.55%
Dividends (1%)	–9.04%	–7.29%	–5.91%	–4.84%
Dividends (2%)	–9.46%	–7.66%	–6.24%	–5.13%
Dividends (3%)	–9.90%	–8.03%	–6.56%	–5.43%
Dividends (4%)	–10.34%	–.841%	–6.90%	–5.73%
Dividends (5%)	–10.80%	–8.79%	–7.24%	–6.04%
Dividends (6%)	–11.26%	–9.18%	–7.58%	–6.35%
Dividends (7%)	–11.74%	–9.58%	–7.93%	–6.67%
Dividends (8%)	–12.22%	–9.99%	–8.29%	–6.99%
Dividends (9%)	–12.71%	–10.40%	–8.65%	–7.31%
Dividends (10%)	–13.22%	–10.81%	–9.01%	–7.64%

FIGURE 14.48 Effects of significant blackouts (different dividend yields and vesting periods).

Figure 14.48 shows the effects of blackouts under different dividend yields and vesting periods, while Figure 14.49 illustrates the results stemming from different dividend yields and suboptimal exercise behavior multiples. Clearly, it is almost impossible to predict the exact impact unless a detailed analysis is performed, but the range can be generalized to be typically between 10 percent and 20 percent. Blackout periods can only be modeled in a binomial lattice and not in the BSM/GBM.

Nonmarketability Issues

The 2004 FAS 123 revision does not explicitly discuss the issue of nonmarketability; that is, ESOs are neither directly transferable to someone else nor freely tradable in the open market. Under such circumstances, it can be argued based on sound financial and economic theory that a nontradable and nonmarketable discount can be appropriately applied to the ESO. However, this is not a simple task.

A simple and direct application of a discount should not be based on an arbitrarily chosen percentage *haircut* on the resulting binomial lattice result. Instead, a more rigorous analysis can be performed using a *put option*. A call option is the contractual right, but not the obligation, to *purchase* the underlying stock at some predetermined contractual strike price within a specified time, while a put option is a contractual right, but not the obligation, to *sell* the underlying stock at some predetermined contractual price within a specified time. Therefore, if the holder of the ESO cannot sell or transfer the

% Difference between no blackout periods versus significant blackouts	Dividends (0%)	Dividends (1%)	Dividends (2%)	Dividends (3%)	Dividends (4%)	Dividends (5%)	Dividends (6%)	Dividends (7%)	Dividends (8%)	Dividends (9%)	Dividends (10%)
Suboptimal Behavior Multiple (1.8)	–1.01%	–1.29%	–1.58%	–1.87%	–2.16%	–2.45%	–2.75%	–3.06%	–3.36%	–3.67%	–3.98%
Suboptimal Behavior Multiple (1.9)	–1.01%	–1.29%	–1.58%	–1.87%	–2.16%	–2.45%	–2.75%	–3.06%	–3.36%	–3.67%	–3.98%
Suboptimal Behavior Multiple (2.0)	–1.87%	–2.29%	–2.72%	–3.15%	–3.59%	–4.04%	–4.50%	–4.96%	–5.42%	–5.90%	–6.38%
Suboptimal Behavior Multiple (2.1)	–1.87%	–2.29%	–2.72%	–3.15%	–3.59%	–4.04%	–4.50%	–4.96%	–5.42%	–5.90%	–6.38%
Suboptimal Behavior Multiple (2.2)	–4.71%	–5.05%	–5.39%	–5.74%	–6.10%	–6.46%	–6.82%	–7.19%	–7.57%	–7.95%	–8.34%
Suboptimal Behavior Multiple (2.3)	–4.71%	–5.05%	–5.39%	–5.74%	–6.10%	–6.46%	–6.82%	–7.19%	–7.57%	–7.95%	–8.34%
Suboptimal Behavior Multiple (2.4)	–4.71%	–5.05%	–5.39%	–5.74%	–6.10%	–6.46%	–6.82%	–7.19%	–7.57%	–7.95%	–8.23%
Suboptimal Behavior Multiple 92.5)	–6.34%	–6.80%	–7.28%	–7.77%	–8.26%	–8.76%	–9.27%	–9.79%	–10.32%	–10.86%	–11.41%
Suboptimal Behavior Multiple (2.7)	–6.34%	–6.80%	–7.28%	–7.77%	–8.26%	–8.76%	–9.27%	–9.79%	–10.32%	–10.86%	–11.41%
Suboptimal Behavior Multiple (2.8)	–6.34%	–6.80%	–7.28%	–7.77%	–8.26%	–8.76%	–9.27%	–9.79%	–10.32%	10.86%	–11.41%
Suboptimal Behavior Multiple (2.9)	–8.62%	–9.04%	–9.46%	–9.9–%	–10.34%	–10.80%	–11.26%	–11.74%	–12.22%	–12.71%	–13.22%
Suboptimal Behavior Multiple (3.0)	–8.62%	–9.04%	–9.46%	–9.90%	–10.34%	–10.80%	–11.26%	–11.74%	–12.22%	–12.71%	–13.22%

FIGURE 14.49 Effects of significant blackouts (different dividend yields and exercise behaviors).

rights of the option to someone else, then the holder of the option has given up his or her rights to a put option (i.e., the employee has written or sold the firm a put option). Calculating the put option and discounting this value from the call option provides a theoretically correct and justifiable nonmarketability and nontransferability discount to the existing option.

However, care should be taken in analyzing this haircut or discounting feature. The same inputs that go into the customized binomial lattice to calculate a call option should also be used to calculate a customized binomial lattice for a put option. That is, the put option must also be under the same risks (volatility that can change over time), economic environment (risk-free rate structure that can change over time), corporate financial policy (a static or changing dividend yield over the life of the option), contractual obligations (vesting, maturity, strike price, and blackout dates), investor irrationality (suboptimal exercise behavior), firm performance (stock price at grant date), and so forth.

Although nonmarketability discounts or haircuts are not explicitly discussed in FAS 123, the valuation analysis is performed here for the sake of completeness. It is up to each firm's management to decide if haircuts should and can be applied. Figure 14.50 shows the customized binomial lattice valuation results of a typical ESO. Figure 14.51 shows the results from a nonmarketability analysis performed using a down-and-in upper barrier modified put option with the same exotic inputs (vesting, blackouts, forfeitures, suboptimal behavior, and so forth) calculated using the customized binomial lattice model.[21] The discounts range from 22 percent to 53 percent. These calculated discounts look somewhat significant but are actually in

Customized Binomial Lattice (Option Valuation)	Behavior (1.20)	Behavior (1.40)	Behavior (1.60)	Behavior (1.80)	Behavior (2.00)	Behavior (2.20)	Behavior (2.40)	Behavior (2.60)	Behavior (2.80)	Behavior (3.00)
Forfeiture (0.00%)	$24.57	$30.53	$36.16	$39.90	$43.15	$45.87	$48.09	$49.33	$50.40	$51.31
Forfeiture (4.00%)	$22.69	$27.65	$32.19	$35.15	$37.67	$39.74	$41.42	$42.34	$43.13	$43.80
Forfeiture (10.00%)	$21.04	$25.22	$28.93	$31.29	$33.27	$34.88	$36.16	$36.86	$37.45	$37.94
Forfeiture (15.00%)	$19.58	$23.13	$26.20	$28.11	$29.69	$30.94	$31.93	$32.46	$32.91	$33.29
Forfeiture (20.00%)	$18.28	$21.32	$23.88	$25.44	$26.71	$27.70	$28.48	$28.89	$29.23	$29.52
Forfeiture (25.00%)	$17.10	$19.73	$21.89	$23.17	$24.20	$25.00	$25.61	$25.93	$26.19	$26.41
Forfeiture (30.00%)	$16.02	$18.31	$20.14	$21.21	$22.06	$22.70	$23.19	$23.44	$23.65	$23.82
Forfeiture (35.00%)	$15.04	$17.04	$18.61	$19.51	$20.20	$20.73	$21.12	$21.32	$21.49	$21.62
Forfeiture (40.00%)	$14.13	$15.89	$17.24	$18.00	$18.58	$19.01	$19.33	$19.49	$19.63	$19.73

FIGURE 14.50 Customized binomial lattice valuation results. (Assumptions used: stock and strike price of $100, 10-year maturity, 1-year vesting, 35% volatility, 0% dividends, 5% risk-free rate, suboptimal exercise behavior multiple range of 1.2–3.0, forfeiture range of 0–40%, and 1,000 step customized lattice.)

Haircut (Customized Binomial Lattice Modified Put)	Behavior (1.20)	Behavior (1.40)	Behavior (1.60)	Behavior (1.80)	Behavior (2.00)	Behavior (2.20)	Behavior (2.40)	Behavior (2.60)	Behavior (2.80)	Behavior (3.00)
Forfeiture (0.00%)	$11.33	$11.33	$11.33	$11.33	$11.33	$11.33	$11.33	$11.33	$11.33	$11.33
Forfeiture (5.00%)	$10.76	$10.76	$10.76	$10.76	$10.76	$10.76	$10.76	$10.76	$10.76	$10.76
Forfeiture (10.00%)	$10.23	$10.23	$10.23	$10.23	$10.23	$10.23	$10.23	$10.23	$10.23	$10.23
Forfeiture (15.00%)	$9.72	$9.72	$9.72	$9.72	$9.72	$9.72	$9.72	$9.72	$9.72	$9.72
Forfeiture (20.00%)	$9.23	$9.23	$9.23	$9.23	$9.23	$9.23	$9.23	$9.23	$9.23	$9.23
Forfeiture (25.00%)	$8.77	$8.77	$8.77	$8.77	$8.77	$8.77	$8.77	$8.77	$8.77	$8.77
Forfeiture (30.00%)	$8.34	$8.34	$8.34	$8.34	$8.34	$8.34	$8.34	$8.34	$8.34	$8.34
Forfeiture (35.00%)	$7.92	$7.92	$7.92	$7.92	$7.92	$7.92	$7.92	$7.92	$7.92	$7.92
Forfeiture (40.00%)	$7.52	$7.52	$7.52	$7.52	$7.52	$7.52	$7.52	$7.52	$7.52	$7.52

Nonmarketability and Nontransferability Discount (%)	Behavior (1.20)	Behavior (1.40)	Behavior (1.60)	Behavior (1.80)	Behavior (2.00)	Behavior (2.20)	Behavior (2.40)	Behavior (2.60)	Behavior (2.80)	Behavior (3.00)
Forfeiture (0.00%)	46.09%	37.09%	31.32%	28.39%	26.25%	24.69%	23.55%	22.96%	22.47%	22.07%
Forfeiture (5.00%)	47.43%	38.92%	33.43%	30.62%	28.57%	27.08%	25.98%	25.42%	24.95%	24.57%
Forfeiture (10.00%)	48.60%	40.55%	35.35%	32.68%	30.73%	29.32%	28.28%	27.75%	27.31%	26.95%
Forfeiture (15.00%)	49.62%	42.01%	37.08%	34.57%	32.73%	31.40%	30.43%	29.93%	29.53%	29.19%
Forfeiture (20.00%)	50.52%	43.31%	38.66%	36.29%	34.57%	33.33%	32.42%	31.96%	31.59%	31.28%
Forfeiture (25.00%)	51.32%	44.48%	40.09%	37.86%	36.25%	35.10%	34.26%	33.84%	33.49%	33.22%
Forfeiture (30.00%)	52.03%	45.53%	41.38%	39.29%	37.79%	36.72%	35.95%	35.56%	35.25%	35.00%
Forfeiture (35.00%)	52.67%	46.48%	42.56%	40.60%	39.20%	38.21%	37.50%	37.15%	36.86%	36.63%
Forfeiture (40.00%)	53.24%	47.34%	43.64%	41.80%	40.49%	39.57%	38.92%	38.60%	38.34%	38.14%

FIGURE 14.51 Nonmarketability and nontransferability discount.

line with market expectations.[22] As these discounts are not explicitly sanctioned by FASB, the author cautions their use in determining the fair-market value of the ESOs.

Expected Life Analysis

As seen previously, the 2004 Final FAS 123 Sections A15 and B64 expressly prohibit the use of a modified BSM with a single expected life. This means that instead of using an expected life as the *input* into the BSM to obtain the similar results as in a customized binomial lattice, the analysis should be done the other way around. That is, using vesting requirements, suboptimal exercise behavior multiples, forfeiture or employee turnover rates, and the other standard option inputs, calculate the valuation results using the customized binomial lattice. This result can then be compared with a modified BSM and the expected life can then be *imputed*. Excel's goal-seek function can be used to obtain the imputed expected life of the option by setting the BSM result equal to the customized binomial lattice. The resulting expected life can then be compared with historical data as a secondary verification of the results, that is, if the expected life falls within reasonable bounds based on historical performance. This is the correct approach because measuring the expected life of an option is very difficult and inaccurate.

Figure 14.52 illustrates the use of Excel's goal-seek function on the ESO Valuation Toolkit software to impute the expected life into the BSM model by setting the BSM results equal to the customized binomial lattice results.

Figure 14.53 illustrates another case where the expected life can be imputed, but this time the forfeiture rates are not set at zero. In this case, the BSM results will need to be modified. For example, the customized binomial lattice result of $5.41 is obtained with a 15 percent forfeiture rate. This means that the BSM result needs to be BSM(1–15%) = $5.41 using the modified expected life method. The expected life that yields the BSM value of $6.36 ($5.41/85% is $6.36, and $6.36(1–15%) is $5.41) is 2.22 years.

Dilution

In most cases, the effects of dilution can be safely ignored as the proportion of ESO grants is relatively small compared to the total equity issued by the company. In investment finance theory, the market has already anticipated

Customized Binomial Lattice Results to Impute the Expected Life for BSM

Applying Different Suboptimal Behavior Multiples

Stock Price	$20.00	$20.00	$20.00	$20.00	$20.00	$20.00	$20.00
Strike Price	$20.00	$20.00	$20.00	$20.00	$20.00	$20.00	$20.00
Maturity	10.00	10.00	10.00	10.00	10.00	10.00	10.00
Risk-Free Rate	3.50%	3.50%	3.50%	3.50%	3.50%	3.50%	3.50%
Dividend	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Volatility	50.00%	50.00%	50.00%	50.00%	50.00%	50.00%	50.00%
Vesting	4.00	4.00	4.00	4.00	4.00	4.00	4.00
Suboptimal Behavior	**1.10**	**1.50**	**2.00**	**2.50**	**3.00**	**3.50**	**4.00**
Forfeiture Rate	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Lattice Steps	1000	1000	1000	1000	1000	1000	1000
Binomial	$8.94	$10.28	$11.03	$11.62	$11.89	$12.18	$12.29
BSM	$12.87	$12.87	$12.87	$12.87	$12.87	$12.87	$12.87
Expected Life	***4.42***	***5.94***	***6.95***	***7.83***	***8.26***	***8.74***	***8.93***
Modified BSM	$8.94	$10.28	$11.03	$11.62	$11.89	$12.18	$12.29

FIGURE 14.52 Imputing the expected life for the BSM using the binomial lattice results.

Customized Binomial Lattice Results to Impute the Expected Life for BSM

Applying Different Forfeiture Rates

Stock Price	$20.00	$20.00	$20.00	$20.00	$20.00	$20.00	$20.00
Strike Price	$20.00	$20.00	$20.00	$20.00	$20.00	$20.00	$20.00
Maturity	10.00	10.00	10.00	10.00	10.00	10.00	10.00
Risk-Free Rate	3.50%	3.50%	3.50%	3.50%	3.50%	3.50%	3.50%
Dividend	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Volatility	50.00%	50.00%	50.00%	50.00%	50.00%	50.00%	50.00%
Vesting	4.00	4.00	4.00	4.00	4.00	4.00	4.00
Suboptimal Behavior	1.50	1.50	1.50	1.50	1.50	1.50	1.50
Forfeiture Rate	**0.00%**	**.250%**	**5.00%**	**7.50%**	**10.00%**	**12.50%**	**15.00%**
Lattice Steps	1000	1000	1000	1000	1000	1000	1000
Binomial	$10.28	$9.23	$8.29	$7.44	$6.69	$6.02	$5.41
BSM	$12.87	$12.87	$12.87	$12.87	$12.87	$12.87	$12.87
Expected Life	***5.94***	***4.71***	***3.77***	***3.03***	***2.45***	***1.99***	***1.61***
Modified BSM*	$10.28	$9.23	$8.29	$7.44	$6.69	$6.02	$5.41
Expected Life	***5.94***	***4.97***	***4.19***	***3.55***	***3.02***	***2.59***	***2.22***
Modified BSM**	$10.28	$9.23	$8.29	$7.44	$6.69	$6.02	$5.41

*Note: Uses the binomial lattice result to impute the expected life for a modified BSM.
**Note: Uses the binomial lattice but also accounts for the Forfeiture rate to modify the BSM.

FIGURE 14.53 Imputing expected life for the BSM using lattice results under nonzero forfeiture rates.

the exercise of these ESOs and the effects have already been accounted for in the stock price. Once a new grant is announced, the stock price will immediately and fully incorporate this news and account for any dilution that may occur. This means that as long as the valuation is performed after the announcement is made, then the effects of dilution are nonexistent. The 2004 FAS 123 revisions do not explicitly provide guidance in this area. Given that FASB provides little guidance on dilution (Section A39), and because forecasting stock prices (as part of estimating the effects of dilution) is fairly difficult and inaccurate at best, plus the fact that the dilution effects are minimal (small in proportion compared to all the equity issued by the firm), the effects of dilution are assumed to be minimal and can be safely ignored.

Applying Monte Carlo Simulation for Statistical Confidence and Precision Control

Next, Monte Carlo simulation can be applied to obtain a range of calculated stock option fair values. That is, any of the inputs into the stock options valuation model can be chosen for Monte Carlo simulation if they are uncertain and stochastic. Distributional assumptions are assigned to these variables, and the resulting option values using the BSM, GBM, path simulation, or binomial lattices are selected as forecast cells. These modeled uncertainties include the probability of forfeiture and the employees' suboptimal exercise behavior.

The results of the simulation are essentially a distribution of the stock option values. Keep in mind that the simulation application here is used to vary the inputs to an options valuation model to obtain a range of results, not to model and calculate the options themselves. However, simulation can be applied both to simulate the inputs to obtain the range of options results and to solve the options model through path-dependent simulation. For instance, the simulated input assumptions are those inputs that are highly uncertain and can vary in the future, such as stock price at grant date, volatility, forfeiture rates, and suboptimal exercise behavior multiples. Clearly, variables that are objectively obtained, such as risk-free rates (U.S. Treasury yields for the next 1 month to 20 years are published), dividend yield (determined from corporate strategy), vesting period, strike price, and blackout periods (determined contractually in the option grant) should not be simulated. In addition, the simulated input assumptions can be correlated. For instance, forfeiture rates can be negatively correlated to stock price—if the firm is doing well, its stock price usually increases, making the option more valuable, thus making the employees less likely to leave and the firm less likely to lay off its employees. Finally, the output forecasts are the option valuation results. In fact, Monte Carlo simulation is allowed and recommended in FAS 123 (Sections B64, B65, and footnotes 48, 52, 74, and 97).

Figure 14.54 shows the results obtained using the customized binomial lattices based on single-point inputs of all the variables. The model takes exotic inputs such as vesting, forfeiture rates, suboptimal exercise behavior multiples, blackout periods, and changing inputs (dividends, risk-free rates, and volatilities) over time. The resulting option value is \$31.42. This analysis can then be extended to include simulation. Figure 14.55 illustrates the use of simulation coupled with customized binomial lattices (Risk Simulator software was used to simulate the input variables).

Rather than randomly deciding on the correct number of trials to run in the simulation, statistical significance and precision control are set up to run the required number of trials automatically. A 99.9 percent statistical confidence on a \$0.01 error precision control was selected and 145,510 simulation trials were run.[23] This highly stringent set of parameters means that an adequate number of trials will be run to ensure that the results will fall within a \$0.01 error variability 99.9 percent of the time. For instance, the simulated average result was \$31.32 (Figure 14.55). This means that 999

Risk-Free Rate		*Volatility*		*Dividend Yield*		*Suboptimal Behavior*	
Year	Rate	Year	Rate	Year	Rate	Year	
1	3.50%	1	35.00%	1	1.00%	1	1.80
2	3.75	2	35.00	2	1.00	2	1.80
3	4.00	3	35.00	3	1.00	3	1.80
4	4.15	4	45.00	4	1.50	4	1.80
5	4.20	5	45.00	5	1.50	5	1.80

		Forfeiture Rate		*Blackout Dates*	
		Year	Rate	Month	Step
Stock Price	$100	1	5.00%	12	12
Strike Price	$100	2	5.00	24	24
Time to Maturity	5	3	5.00	36	36
Vesting Period	1	4	5.00	48	48
Lattice Steps	60	5	5.00	60	60
Option value	**$31.42**				

FIGURE 14.54 Single-point result using a customized binomial lattice.

Statistic	Value	Precision
Trials	145,510	
Mean	$31.32	$0.01
Median	$31.43	$0.02
Mode	—	
Standard Deviation	$1.57	$0.01
Variance	$2.46	
Skewness	−0.21	
Kurtosis	2.43	
Coeff. Of Variability	0.05	
Range Minimum	$26.59	
Range Maximum	$35.62	
Range Width	$9.03	
Mean Std. Error	$0.00	

*Tested for $0.01 precision at 99.90% confidence.

FIGURE 14.55 Options valuation result at $0.01 precision with 99.9 percent confidence.

out of 1,000 times, the true option value will be accurate to within $0.01 of $31.32. These measures are statistically valid and objective.[24]

Number of Steps

The higher the number of lattice steps, the higher the precision of the results. Figure 14.56 illustrates the convergence of results obtained using a BSM closed-form model on a European call option without dividends, and comparing its results to the basic binomial lattice. Convergence is generally achieved at 1,000 steps. As such, the analysis results will use 1,000 steps

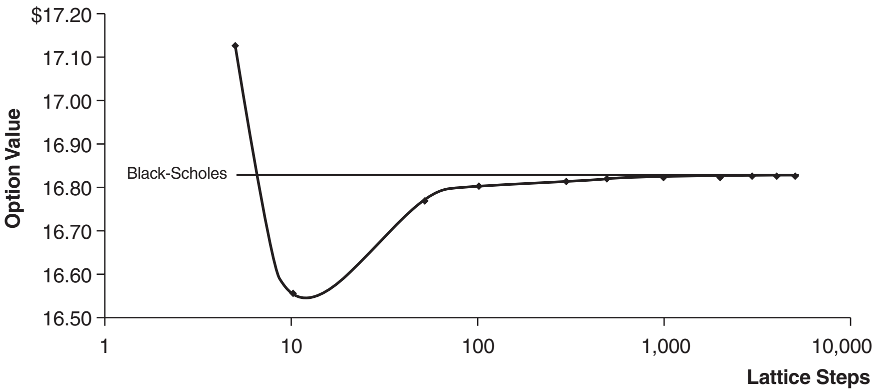


FIGURE 14.56 Convergence of the binomial lattice to closed-form solutions.

whenever possible.[25] Due to the high number of steps required to generate the results, software-based mathematical algorithms are used.[26] For instance, a nonrecombining binomial lattice with 1,000 steps has a total of 2×10^{301} nodal calculations to perform, making manual computation impossible without the use of specialized algorithms.[27] Figure 14.57 illustrates the calculation of convergence by using progressively higher lattice steps. The progression is based on sets of 120 steps (12 months per year multiplied by 10 years). The results are tabulated and the median of the average results is calculated. It shows that 4,200 steps is the best estimate in this customized binomial lattice, and this input is used throughout the analysis.[28]

Conclusion

It has been more than 30 years since Fisher Black, Myron Scholes, and Robert Merton derived their option pricing model and significant advancements have been made; therefore, do not restrict stock option pricing to one specific model (the BSM/GBM) while a plethora of other models and applications can be explored. The three mainstream approaches to valuing stock options are closed-form models (e.g., BSM, GBM, and American option approximation models), Monte Carlo simulation, and binomial lattices. The BSM and GBM will typically *overstate* the fair value of ESOs where there is suboptimal early exercise behavior coupled with vesting requirements and option forfeitures. In fact, firms using the BSM and GBM to value and expense ESOs may be *significantly* overstating their true expense. The BSM requires many underlying assumptions before it works and, as such, has significant limitations, including being applicable only for European options without dividends. In addition, American option approximation models are

Stock Price	$45.17	$45.17	$45.17	$45.17	$45.17	$45.17	$45.17	$45.17	$45.17	$45.17	$45.17	$45.17	$45.17	$45.17
Strike Price	$45.17	$45.17	$45.17	$45.17	$45.17	$45.17	$45.17	$45.17	$45.17	$45.17	$45.17	$45.17	$45.17	$45.17
Maturity	10	10	10	10	10	10	10	10	10	10	10	10	10	10
Risk-Free Rate	1.21%	1.21%	1.21%	1.21%	1.21%	1.21%	1.21%	1.21%	1.21%	1.21%	1.21%	1.21%	1.21%	1.21%
Volatility	49.91%	49.91%	49.91%	49.91%	49.91%	49.91%	49.91%	49.91%	49.91%	49.91%	49.941%	49.91%	49.91%	49.91%
Dividend	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Lattice Steps	10	50	100	120	600	1200	1800	2400	3000	3600	4200	4800	5400	6000
Suboptimal Behavior	1.8531	1.8531	1.8531	1.8531	1.8531	1.8531	1.8531	1.8531	1.8531	1.8531	1.8531	1.8531	1.8531	1.8531
Vesting	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08
Binomial Option Value	$20.55	$17.82	$17.32	$18.55	$17.55	$13.08	$13.11	$12.93	$12.88	$12.91	$13.00	$13.08	$12.93	$13.06

Segments	Steps	Results	Average
1	120	$18.55	$13.91
5	600	$17.55	$13.48
10	1200	$13.08	$13.00
15	1800	$13.11	$12.99
20	2400	$12.93	$12.97
25	3000	$12.83	$12.97
30	3600	$12.91	$12.99
35	4200	$13.00	$13.00
40	4800	$13.03	$13.02
45	5400	$12.93	$12.99
50	6000	$13.06	$13.06
		Median	$13.00

FIGURE 14.57 Convergence of the customized binomial lattice.

very complex and difficult to create in a spreadsheet. The BSM *cannot* account for American options, options based on stocks that pay dividends (the GBM model can, however, account for dividends in a European option), forfeitures, underperformance, stock price barriers, vesting periods, changing business environments and volatilities, suboptimal early exercise behavior, and a slew of other conditions. Monte Carlo simulation when used alone is another option valuation approach, but is restricted only to European options. Simulation can be used in two different ways: to solve the option's fair-market value through path simulations of stock prices, or used in conjunction with other approaches (e.g., binomial lattices and closed-form models) to capture multiple sources of uncertainty in the model.

Binomial lattices are flexible and easy to implement. They are capable of valuing American-type stock options with dividends but require computational power. Software applications should be used to facilitate this computation. Binomial lattices can be used to calculate American options paying dividends and can be easily adapted to solve ESOs with exotic inputs and used in conjunction with Monte Carlo simulation to account for the uncertain input assumptions (e.g., probabilities of forfeiture, suboptimal exercise behavior, vesting, underperformance) and to obtain a high precision at statistically valid confidence intervals. Based on the analyses throughout the case study, it is recommended that the use of a model that assumes an ESO is European style when, in fact, the option is American style with the other exotic variables should not be permitted, as this substantially overstates compensation expense. Many factors influence the fair-market value of ESOs, and a binomial lattice approach to valuation that considers these factors should be used. With due diligence, real-life ESOs can absolutely be valued using the customized binomial lattice approach as shown in this case study, where the methodology employed is pragmatic, accurate, and theoretically sound.

PART Nine

Risk Management

CHAPTER 15

The Warning Signs

The finding of absence is very different from an absence of findings. How does management appropriately evaluate the validity and applicability of analytical results? How should management challenge the assumptions used in the analysis? What are some of the questions that should be asked? This chapter deals with some of the more difficult questions when evaluating the results of Monte Carlo simulation, time-series forecasting, stochastic optimization, and real options analysis.

THE PROBLEM OF NEGLIGENT ENTRUSTMENT

Power tools such as Risk Simulator and Real Options Super Lattice Solver took years to build and many more years to be perfected. It is extremely likely that a new user can simply pick up software products such as these and hit the ground running immediately. However, some knowledge of the theoretical underpinnings is required. In short, to create and perform sophisticated modeling, the analyst first needs to understand some of the underlying assumptions and approaches used in these analytics. Otherwise, it is akin to giving a 3-year-old child a loaded machine gun. The correct term for this situation might be "negligent entrustment." In fact, when the rubber meets the road, more often than not, even so-called *power users* are perplexed and have a difficult time using these tools with respect to their models and business cases. These software tools, despite their analytical power, are just tools. They do not replace the analyst in any way. In fact, tools such as these only accouter the analyst with the appropriate analytics at their fingertips and do not by themselves make the relevant decisions. Such tools only relieve the analyst from having facility with fancy mathematics in order to build sophisticated models. As stated previously, 50 percent of the challenge in decision making is simply thinking about the problem, with 25 percent being the actual modeling and analytics, and the remaining 25 percent being able to convince and explain the results to senior management, clients, colleagues, and yourself.

MANAGEMENT'S DUE DILIGENCE

It might be the job of the analyst to create the models and use the fancy analytics, but it is senior management's job to challenge the assumptions and results obtained from said analysis. For instance, Figure 15.1 lists some of the issues that may arise when running a multivariate regression analysis and time-series forecasting. Although it may not be senior management's job to understand the mathematical or theoretical implications of these issues, management must nonetheless have a good grasp of what they mean.

The following sections are written specifically for senior management who are recipients of different types of advanced analyses results. The next section starts off with a general set of warning signs and moves on to the specifics of each analytical methodology used throughout this book.

SINS OF AN ANALYST

In general, warning signs can be grouped into five categories:

1. Model errors.
2. Assumption and input errors.
3. Analytical errors.
4. User errors.
5. Interpretation errors.

- Out of Range Forecasts
- Structural Breaks
- Specification Errors
- Omitted and Redundant Variables
- Heteroskedasticity and Homoskedasticity
- Multicollinearity
- Spurious Regression and Time Dependency
- Autocorrelation and Serial Correlation
- Correlation versus Causation
- Random Walks
- Mean Reversions
- Jump Processes
- Stochastic Processes

FIGURE 15.1 Warning signs in regression analysis.

Model errors are the errors an analyst would make while creating models. For instance, a financial model created in Excel may have errors stemming from broken links, incorrect functions and equations, poor modeling practices, or a break in the knowledge transfer between the originator of the model and subsequent users as well as successors of the model. This error can be eliminated through diligence on the part of the model creator. Good model-building practices also can assist in eliminating messy models. These practices include:

- Good documentation of the approaches used in the model as well as the integration and connectivity of the subparts that exist in the model.
- Creating a starting page that is linked through hyperlinks or macros with sufficient descriptions of each subpage or worksheet.
- Differentiating assumption input sheets from the models actually performing the number crunching, and from the results or reports page.
- Allowing changes to be made only on the input assumptions page and not directly in the model to prevent accidentally breaking the model.

For a detailed listing of good model-building practices and modeling etiquette, refer to Chapter 3, A Guide to Model-Building Etiquette.

Assumption and input errors are more difficult to tackle. These errors include the inputs required to make the model compute; for example, items such as levels of competitive threats, levels of technological success, revenue projections, income growth rates, market share determination, and so forth. Many of these determinant factors are almost impossible to determine. In fact, the old adage of *garbage in, garbage out* holds true here. The analyst can only do so much.

Multiple approaches exist to help clean up these so-called garbage assumptions. One way is simply to use expert knowledge and advice. For instance, the Delphi method requires the presence of a group of expert engineers in a room to discuss the levels of technological success rates. These engineers with intimate knowledge of the potential success rates are able to provide valuable insights that would otherwise be unavailable to a financial analyst sitting in front of a computer, far removed from the everyday technological challenges. Senior management, based on their many years of experience and expertise, can often provide valuable insights into what certain market outcomes may be. A double-blind experiment also can be conducted, where experts in a group are asked on anonymous questionnaires what their objective estimates of an outcome are. These quantitative outcomes are then tabulated and, on occasion, more experienced participants' comments will be weighted more heavily. The expected value is then used in the model. Here, Monte Carlo simulation can be applied on the distribution of the outcomes related to these expert testimonies. A custom distribution can be constructed using Risk Simulator, which relates back to the weights given to each outcome, or a simple nonparametric custom distribution simulation can also be applied on all possible outcomes obtained. Obviously, if there are ample historical data, then it is relatively easier to project the future, whether it is using some time-series forecast, regression analysis, or Monte Carlo simulation. *When in doubt, simulate!* Instead of arguing and relying on a particular single-point input value of a particular variable, an analyst can just simulate it around the potential outcomes of that input, whether it is the worst-case scenario, nominal-case scenario, or best-case scenario using a triangular distribution or some other distribution through expert assumptions.

No matter the approach used to obtain the data, management must test and challenge these assumptions. One way is to create tornado and sensitivity charts. The variables that drive the bottom line the most (the variable of interest, e.g., net present value, net income, return on investment) that are unpredictable and subject to uncertain levels of fluctuations are the ones that management should focus on. These critical success factors are the ones that management should care about, not some random variable that has little to no effect on the bottom line no matter how attractive or important the variable may be in other instances.

The upshot being that the more expert knowledge and historical data that exist, the better the assumption estimates will be. A good test of the assumptions used is through the application of back-casting, as opposed to forecasting, which looks forward into the future. Back-casting uses historical data to test the validity of the assumptions. One approach is to take the historical data, fit them to a distribution using Risk Simulator's distributional-fitting routines, and test the assumption input. Observe where the assumption value falls within this historical distribution. If it falls outside of the distribution's normal set of parameters (e.g., 95 percent or 99 percent confidence intervals), then the analyst should be able to better describe why there will be a potential structural shift going forward (e.g., mergers and acquisition, divestiture, reallocation of resources, economic downturn, entry of formidable competition, and so forth). In forecasting, similar approaches can be used such as historical data-fitting of the forecast model and holdout approaches (i.e., some historical data are left out in the original forecast model but are used in the subsequent forecast-fitting to verify the model's accuracy).

READING THE WARNING SIGNS IN MONTE CARLO SIMULATION

Monte Carlo simulation is a very potent methodology. Statisticians and mathematicians sometimes dislike it because it solves difficult and often intractable problems with too much simplicity and ease. Instead, mathematical purists would prefer the more elegant approach: the old-fashioned way. Solving a fancy stochastic mathematical model provides a sense of accomplishment and completion as opposed to the brute force method. Monte Carlo creates artificial futures by generating thousands and even millions of sample paths of outcomes and looks at their prevalent characteristics. For analysts in a company, taking graduate-level advanced mathematics courses is neither logical nor practical. A brilliant analyst would use all available tools at his or her disposal to obtain the same answer the easiest way possible. One such tool is Monte Carlo simulation using Risk Simulator. The major benefit that Risk Simulator brings is its simplicity of use. The major downfall that Risk Simulator brings is also its simplicity of use!

Here are 14 due-diligence issues management should evaluate when an analyst presents a report with a series of advanced analytics using simulation.

1. How Are the Distributions Obtained? One thing is certain: If an analyst provides a report showing all the fancy analyses undertaken and one of these analyses is the application of Monte Carlo simulation of a few dozen variables, where each variable has the same distribution (e.g., triangular distribution), management should be very worried indeed and with good reason. One might be able to accept the fact that a few variables are triangularly distributed, but to assume that this holds true for several dozen other variables is ludicrous. One way to test the validity of distributional assumptions is to apply historical data to the distribution and see how far off one is.

Another approach is to take the distribution and test its alternate parameters. For instance, if a normal distribution is used on simulating market share, and the mean is set at 55 percent with a standard deviation of 45 percent, one should be extremely worried. Using Risk Simulator's alternate-parameter function, the 10th and 90th percentiles indicate a value of –2.67 percent and 112.67 percent. Clearly these values cannot exist under actual conditions. How can a product have –2.67 or 112.67 percent of the market share? The alternate-parameters function is a very powerful tool to use in conditions such as these. Almost always, the first thing that should be done is the use of alternate parameters to ascertain the logical upper and lower values of an input parameter.

2. How Sensitive Are the Distributional Assumptions? Obviously, not all variables under the sun should be simulated. For instance, a U.S.-based firm doing business within the 48 contiguous states should not have to worry about what happens to the foreign exchange market of the Zairian zaire. Risk is something one bears and is the outcome of uncertainty. Just because there is uncertainty, there could very well be no risk. If the only thing that bothers a U.S.-based firm's CEO is the fluctuation of the Zairian zaire, then

I might suggest shorting some zaires and shifting his or her portfolio to U.S.-based bonds.

In short, simulate when in doubt, but simulate the variables that actually have an impact on what you are trying to estimate. Two very powerful tools to decide which variables to analyze are tornado and sensitivity charts. Make sure the simulated variables are the critical success factors—variables that have a significant impact on the bottom line being estimated while at the same time being highly uncertain and beyond the control of management.

3. What Are the Critical Success Factors? Critical success factors are related to how sensitive the resulting bottom line is to the input variables and assumptions. The first step that should be performed before using Monte Carlo simulation is the application of tornado charts. Tornado charts help identify which variables are the most critical to analyze. Coupled with management's and the analyst's expertise, the relevant critical success factors—the variables that drive the bottom line the most while being highly uncertain and beyond the control of management—can be determined and simulated. Obviously the most sensitive variables should receive the most amount of attention.

4. Are the Assumptions Related, and Have Their Relationships Been Considered? Simply defining assumptions on variables that have significant impact without regard to their interrelationships is also a major error most analysts make. For instance, when an analyst simulates revenues, he or she could conceivably break the revenue figures into price and quantity, where the resulting revenue figure is simply the product of price and quantity. The problem is if both price and quantity are considered as independent variables occurring in isolation, a major error arises. Clearly, for most products, the law of demand in economics takes over, where the higher the price of a product, ceteris paribus, or holding everything else constant, the quantity demanded of the same product decreases. Ignoring this simple economic truth, where both price and quantity are assumed to occur independently of one another, means that the possibility of a high price and a high quantity demanded may occur simultaneously, or vice versa. Clearly this condition will never occur in real life; thus, the simulation results will most certainly be flawed. The revenue or price estimates can also be further disaggregated into several product categories, where each category is correlated to the rest of the group (competitive products, product life cycle, product substitutes, complements, and cannibalization). Other examples include the possibility of economies of scale (where a higher production level forces cost to decrease over time), product life cycles (sales tend to decrease over time and plateau at a saturation rate), and average total costs (the average of fully allocated cost decreases initially and increases after it hits some levels of diminishing returns). Therefore, relationships, correlations, and causalities have to be modeled appropriately. If data are available, a simple correlation matrix can be generated through Excel to capture these relationships.

5. Considered Truncation? Truncation is a major error Risk Simulator users commit, especially when using the infamous triangular distribution. The triangular distribution is very simple and intuitive. As a matter of fact, it is probably the most widely used distribution in Risk Simulator, apart from the normal and uniform distributions. Simplistically, the triangular distribution looks at the minimum value, the most probable value, and the maximum value. These three inputs are often confused with the worst-case, nominal-case, and best-case scenarios. This assumption is indeed incorrect.

In fact, a worst-case scenario can be translated as a highly unlikely condition that *will* still occur given a percentage of the time. For instance, one can model the economy as high, average, and low, analogous to the worst-case, nominal-case, and best-case scenarios. Thus, logic would dictate that the worst-case scenario might have, say, a 15 percent chance of occurrence, the nominal-case a 50 percent chance of occurrence, and a 35 percent chance that a best-case scenario will occur. This approach is what is meant by using a best-, nominal-, and worst-case scenario analysis. However, compare that to the triangular distribution, where the minimum and maximum cases will almost never occur, with a probability of occurrence set at zero!

For instance, see Figure 15.2, where the worst-, nominal-, and best-case scenarios are set as 5, 10, and 15, respectively. Note that at the extreme values, the probability of 5 or 15 occurring is virtually zero, as the areas under the curve (the measure of probability) of these extreme points are zero. In other words, 5 and 15 will almost *never* occur. Compare that to the economic scenario where these extreme values have either a 15 percent or 35 percent chance of occurrence. Instead, distributional truncation should be considered here. The same applies to any other distribution. Figure 15.3 illustrates a truncated normal distribution where the extreme values do not extend to both positive and negative infinities, but are truncated at 7 and 13.

6. How Wide Are the Forecast Results? I have seen models that are as large as 30MB with over 1,000 distributional assumptions. When you have a model that big with so many assumptions, there is a huge problem! For one, it takes an unnecessarily long time to run in Excel, and for another, the results generated are totally bogus. One thing is certain: The final forecast distribution of the results will most certainly be too large to use to make any definitive decision. Besides, what is the use of generating results that are close to a range between negative and positive infinity?

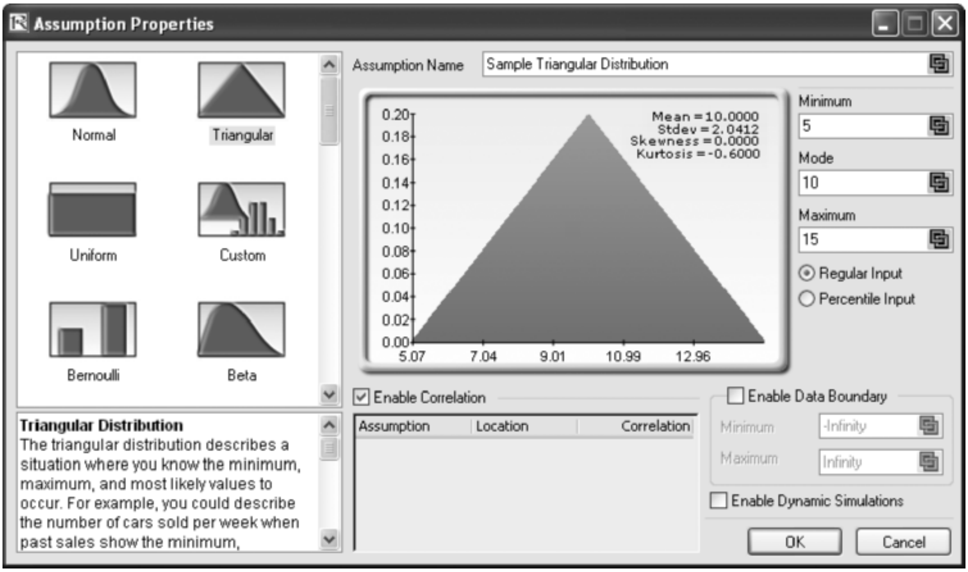


FIGURE 15.2 Sample triangular distribution.

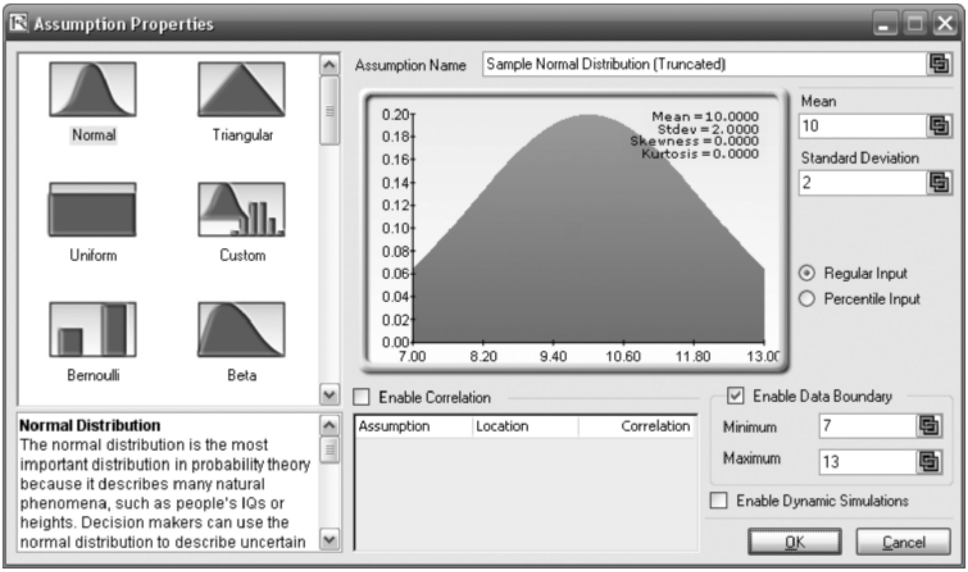


FIGURE 15.3 Truncating a distribution.

The results that you obtain should fall within decent parameters and intervals. One good check is to simply look at the single-point estimates. In theory, the single-point estimate is based on all precedent variables at their respective expected values. Thus, if one perturbs these expected values by instituting distributions about their single-point estimates, then the resulting single-point bottom line estimate should also fall within this forecast interval.

7. What Are the End Points and Extreme Values? Mistaking end points is both an error of interpretation and a user error. For instance, Figure 15.4 illustrates the results obtained from a financial analysis with extreme values between \$5.49 million and \$15.49 million. By making the leap that the worst possible outcome is \$5.49 million and the best possible outcome is \$15.49 million, the analyst has made a major error. Clicking on the *Options* menu and *Data Filter* area, one can choose any display range (Figure 15.5).

Clearly, if the *show data less than 2 standard deviations* option is chosen, the graph looks somewhat different (the endpoints are now 6.34 and 14.34 as compared to 5.49 and 15.49), indicating the actual worst and best cases (Figure 15.6). Of course, the interpretation would be quite different here than with the 2.0 standard deviations option chosen.

8. Are There Breaks Given Business Logic and Business Conditions? Assumptions used in the simulation may be based on valid historical data, which means that the distributional outcomes would be valid if the firm indeed existed in the past. However, going forward, historical data may not be the

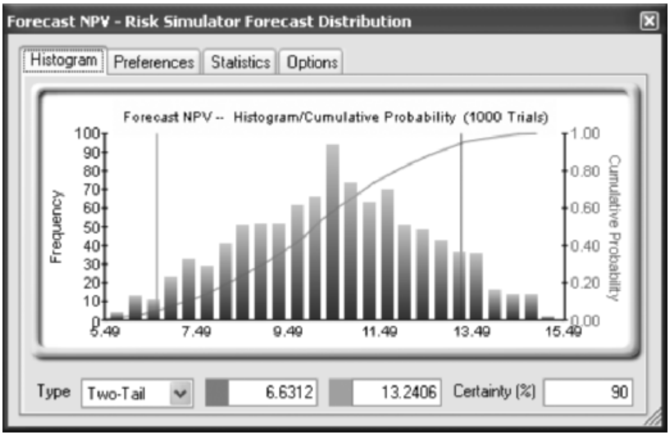


FIGURE 15.4 Truncated extreme values.

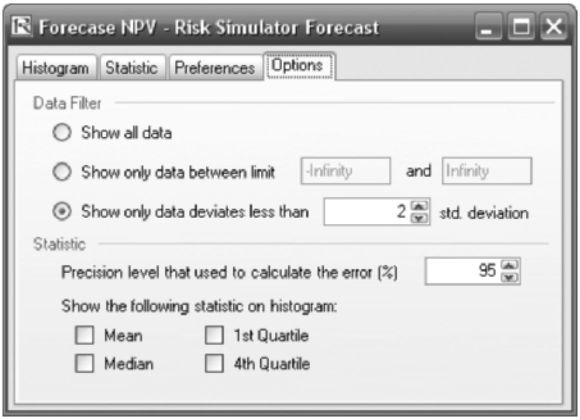


FIGURE 15.5 Display range preferences.

best predictor of the future. In fact, past performance is no indicator of future ability to perform, especially when structural breaks in business conditions are predicted to occur. Structural breaks include situations where firms decide to go global, acquire other firms, divest part of their assets, enter into new markets, and so forth. The resulting distributional forecasts need to be revalidated based on these conditions. The results based on past performance could be deemed as the base-case scenario, with additional adjustments and add-ons as required. This situation is especially true in the research and development arena, where by definition of research and development,

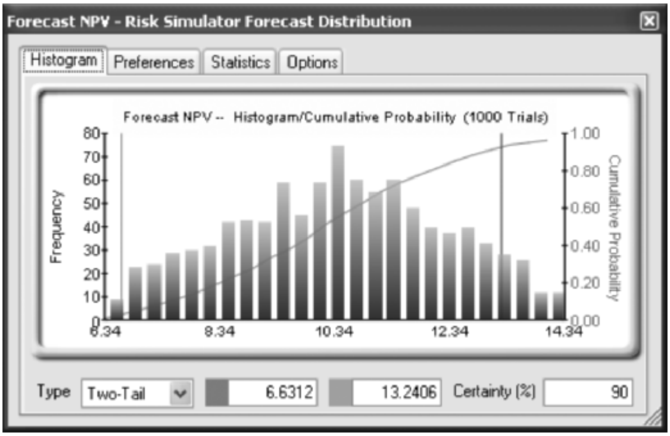


FIGURE 15.6 Display range using fixed end points.

things that are yet to be developed are new and novel in nature; thus by definition, there exist no historical data on which to base the future forecasts. In situations such as these, it is best to rely on experience and expert opinions of future outcomes. Other approaches where historical data do not exist include using market proxies and project comparables—where current or historical projects and firms with similar functions, markets, and risks are used as benchmarks.

9. Do the Results Fall Within Expected Economic Conditions? One of the most dangerous traps analysts fall into is the trap of data mining. Rather than relying on solid theoretical frameworks, analysts let the data sort things out by themselves. For instance, analysts who blindly use stepwise regression and distributional fitting fall directly into this data-mining trap. Instead of relying on theory a priori, or before the fact, analysts use the results to explain the way things look, a posteriori, or after the fact.

A simple example is the prediction of the stock market. Using tons of available historical data on the returns of the Standard & Poor's 500 index, an analyst runs a multivariate stepwise regression using over a hundred different variables ranging from economic growth, gross domestic product, and inflation rates, to the fluctuations of the Zairian zaire, to who won the Super Bowl and the frequency of sunspots on particular days. Because the stock market by itself is unpredictable and random in nature, as are sunspots, there seems to be some relationship over time. Although this relationship is purely spurious and occurred out of happenstance, a stepwise regression and correlation matrix will still pick up this spurious relationship and register the relationship as statistically significant. The resulting analysis will show that sunspots do in fact explain fluctuations in the stock market. Therefore, is the analyst correct in setting up distributional assumptions based on sunspot activity in the hopes of beating the market? When one throws a computer at data, it is almost certain that a spurious connection will emerge.

The lesson learned here is to look at particular models with care when trying to find relationships that may seem on the surface to be valid, but in fact are spurious and accidental in nature, and that holding all else constant, the relationship dissipates over time. Merely correlating two randomly occurring events and seeing a relationship is nonsense and the results should not be accepted. Instead, analysis should be based on economic and financial rationale. In this case, the economic rationale is that the relationship between sunspots and the stock market are completely accidental and should thus be treated as such.

10. What Are the Values at Risk? Remember the story about my friend and me going skydiving in the first chapter? Albeit fictitious, it illustrates the differences between risk and uncertainty. When applying Monte Carlo simulation, an analyst is looking at uncertainty; that is, distributions are applied to different variables that drive a bottom-line forecast. Figure 15.7 shows a very simple calculation, where on a deterministic basis, if revenue is \$2, cost is \$1, the resulting net income is simply \$1 (i.e., \$2 – \$1). However, in the dynamic model, where revenue is “around \$2,” cost is “around \$1,” the net income is “around \$1.” This “around” comment signifies the uncertainty involved in each of these variables. The resulting variable will also be an “around” number. In fact, when Risk Simulator is applied, the resulting single-point estimate also ends up being \$1. The only difference being that there is a forecast distribution surrounding this \$1 value. By performing Monte Carlo simulation, a level of uncertainty surrounding this single-point estimate is obtained. Risk analysis has *not yet* been done. Only uncertainty analysis has been done thus far. By running simulations, only the levels of uncertainty have been quantified if the reports are shown but the results are not used to adjust for risk.

For instance, one can in theory simulate everything under the sun, including the fluctuations of the Zairian zaire, and if the Zairian zaire has no impact on the project being analyzed, capturing the uncertainty surrounding the currency does not mean one has managed, reduced, or analyzed the project’s foreign exchange risks. It is only when the results are analyzed and used appropriately that risk analysis has been done. For instance, if an

STATIC MODEL			DYNAMIC MODEL			
Revenue	\$	2.00	Revenue	\$	2.00	<<--- This is an Input Assumption
Cost	\$	1.00	Cost	\$	1.00	<<--- This is an Input Assumption
Income	\$	1.00	Income	\$	1.00	<<--- This is an Output Forecast

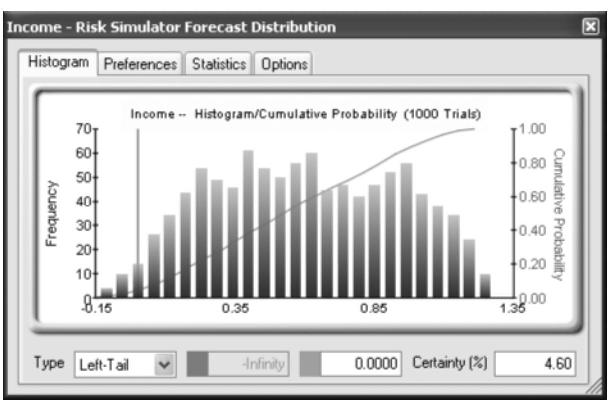


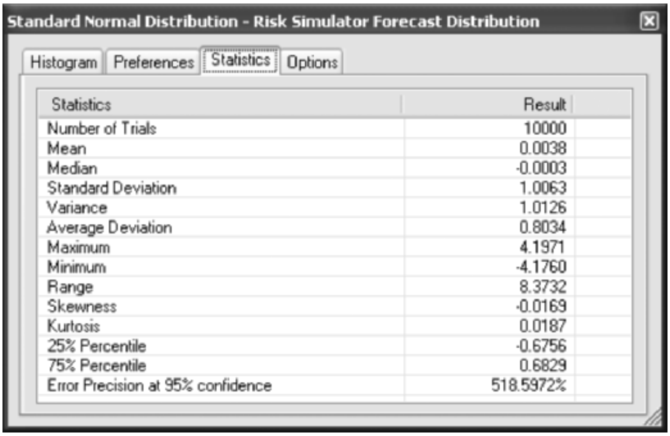
FIGURE 15.7 Illustrating the differences between risk and uncertainty.

analyst is evaluating three similar projects where each project has an expected value of \$1 in net income but with different distributions, no new information is realized. However, when the results are used appropriately, where we say the first project has a \$0.30 value at the fifth percentile, while the second and third projects have \$0.20 and –\$0.10 values at the fifth percentile, will risk analysis have been done. Holding everything else constant, the best project is clearly the first project, where in the worst-case scenario 5 percent of the time, the minimum amount to be gained is \$0.30, the largest of the three. Obviously, other measures can be used, including the mean divided by the standard deviation (creating a coefficient of variability or bang-for-the-buck measure), risk-adjusted return on capital (RAROC or median less the fifth percentile divided by the volatility), and so forth, as detailed in Chapter 2, From Risk to Riches. Suffice it to say, as long as the risk adjustment is applied appropriately across all projects for comparability purposes, the measurement will be valid. The upshot being that simply noting the uncertainty levels around a value is not risk analysis. It is only when this value is adjusted according to its risk levels has risk analysis actually been performed.

11. How Do the Assumptions Compare to Historical Data and Knowledge? Suspect distributional assumptions should be tested through the use of back-casting, which uses historical data to test the validity of the assumptions. One approach is to take the historical data, fit them to a distribution using Risk Simulator's distributional-fitting routines, and test the assumption inputs. See if the distributional-assumption values fall within this historical distribution. If they fall outside of the distribution's normal set of parameters (e.g., 95 percent or 99 percent confidence intervals), then the analyst should better be able to describe and explain this apparent discontinuity, which can very well be because of changing business conditions and so forth.

12. How Do the Results Compare Against Traditional Analysis? A very simple test of the analysis results is through single-point estimates. For instance, remember the \$1 net income example? If the single-point estimate shows \$1 as the expected value of net income, then, in theory, the uncertainty surrounding this \$1 should have the initial single-point estimate somewhere within its forecast distribution. If \$1 is not within the resulting forecast distribution, something is amiss here. Either the model used to calculate the original \$1 single-point estimate is flawed or the simulation assumptions are flawed. To recap, how can "around \$2" minus "around \$1" not be "around \$1"?

13. Do the Statistics Confirm the Results? Risk Simulator provides a wealth of statistics after performing a simulation. Figure 15.8 shows a sample listing



Statistics	Result
Number of Trials	10000
Mean	0.0038
Median	-0.0003
Standard Deviation	1.0063
Variance	1.0126
Average Deviation	0.8034
Maximum	4.1971
Minimum	-4.1760
Range	8.3732
Skewness	-0.0169
Kurtosis	0.0187
25% Percentile	-0.6756
75% Percentile	0.6829
Error Precision at 95% confidence	518.5972%

FIGURE 15.8 Standard-normal distribution statistics.

of these statistics, which can be obtained through the *View* | *Statistics* menu in Risk Simulator. Some of these statistics when used in combination provide a solid foundation of the validity of the results. When in doubt as to what the *normal*-looking statistics should be, simply run a simulation in Risk Simulator and set the distribution to normal with a mean of 0.00 and a standard deviation of 1.00. This condition would create a standard-normal distribution, one of the most basic statistical distributions. The resulting set of statistics is shown in Figure 15.8. See Chapter 2, From Risk to Riches, for more details on some basic statistics and interpreting distributional moments.

Clearly, after running 10,000 trials, the resulting mean is 0.00 with a standard deviation of 1.00, as specified in the assumption. Of particular interest are the skewness and kurtosis values. For a normally distributed result, the skewness is close to 0.00, and the excess kurtosis is close to 0.00. If the results from your analysis fall within these parameters, it is clear that the forecast values are symmetrically distributed with no excess areas in the tail. A highly positive or negative skew would indicate that something might be going on in terms of some distributional assumptions that are skewing the results either to the left or to the right. This skew may be intentional or something is amiss in terms of setting up the relevant distributions. Also, a significantly higher kurtosis value would indicate that there is a higher probability of occurrence in the tails of the distribution, which means extreme values or catastrophic events are prone to occur more frequently than predicted in most normal circumstances. This result may be expected or not. If not, then the distributional assumptions in the model should be revisited with greater care, especially the extreme values of the inputs.

14. Are the Correct Methodologies Applied? The problem of whether the correct methodology is applied is where user error comes in. The analyst should be able to clearly justify why a lognormal distribution is used instead of a uniform distribution, and so forth, and why distributional fitting is used instead of bootstrap simulation, or why a tornado chart is used instead of a sensitivity chart. All of these methodologies and approaches require some basic levels of understanding, and questions such as these are most certainly required as part of management's due diligence when evaluating an analyst's results.

> Warning signs to watch out for in Monte Carlo simulation and questions to ask include how the distributions are obtained, how sensitive are the distributional assumptions, how to identify the critical success factors, how the distributional assumptions are related, if the distributions are truncated, how wide are the forecast values, what are the end points and extreme values, are there breaks in business logic and conditions, do the results follow economic rationale, what are the values-at-risk, how do the results compare with historical data and knowledge, how do the results compare with traditional analyses, do the statistics confirm expectations, and are the correct methodologies applied.

READING THE WARNING SIGNS IN TIME-SERIES FORECASTING AND REGRESSION

Another frequently used decision-analysis tool is forecasting. One thing is certain: You can never predict the future with perfect accuracy. The best that you can hope for is to get as close as possible. In addition, it is actually okay to be wrong on occasion. As a matter of fact, it is sometimes good to be wrong, as valuable lessons can be learned along the way. It is better to be wrong consistently than to be wrong on occasion, because if you are wrong consistently in one direction, you can correct or reduce your expectations, or increase your expectations when you are consistently overoptimistic or underoptimistic. The problem arises when you are occasionally right and occasionally wrong, and you have no idea when or why it happens. Some of the issues that should be addressed when evaluating time-series or any other forecasting results include the following:

1. Out of Range Forecasts Not all variables can be forecast using historical data. For instance, did you know that you can predict, rather reliably, the ambient temperature given the frequency of cricket chirps? Collect a bunch of crickets and change the ambient temperature, collect the data, and run a bivariate regression, and you would get a high level of confidence as seen in the coefficient of determination or R-squared value. Given this model, you could reasonably predict ambient temperature whenever crickets chirp, correct? Well, if you answered yes, you have just fallen into the trap of forecasting out of range.

Suppose your model holds up to statistical scrutiny, which it may very well do, assuming you do a good job with the experiment and data collection. Using the model, one finds that crickets chirp more frequently the higher the ambient temperature, and less frequently the colder it gets. What do you presume would happen if one were to toss a poor cricket in the oven and turn it up to 550 degrees? What happens when the cricket is thrown into the freezer instead? What would occur if a Malaysian cricket were used instead of the Arizona reticulated cricket? The quick answer is you can toss your fancy statistical regression model out the window if any of these things happened. As for the cricket in the oven, you would most probably hear the poor thing give out a very loud chirp and then complete silence. Regression and prediction models out of sample, that is, modeling events that are out of place and out of the range of the data collected in ordinary circumstances, on occasion will fail to work, as is clearly evident from the poor cricket.

2. Structural Breaks Structural breaks in business conditions occur all the time. Some example instances include going public, going private, merger, acquisition, geographical expansion, adding new distribution channels, existence of new competitive threats, union strikes, change of senior management, change of company vision and long-term strategy, economic downturn, and so forth. Suppose you are an analyst at FedEx performing volume, revenue, and profitability metric forecasting of multiple break-bulk stations. These stations are located all around the United States and each station has its own seasonality factors complete with detailed historical data. Some advanced econometric models are applied, ranging from ARIMA (autoregressive integrated moving average) and ECM (error correction models) to GARCH (generalized autoregressive conditional heteroskedasticity) models; these time-series forecasting models usually provide relatively robust forecasts. However, within a single year, management reorganization, union strikes, pilot strikes, competitive threats (UPS, your main competitor, decided to enter a new submarket), revised accounting rules, and a plethora of other *coincidences* simply made all the forecasts invalid. The analyst must decide if these coincidences are just that, coincidences, or if they point to a fundamental structural change in the way global freight businesses are run. Obviously, certain incidences are planned or expected, whereas others are unplanned and unexpected. The planned incidences should thus be considered when performing forecasting.

3. Specification Errors Sometimes, models are incorrectly specified. A nonlinear relationship can be very easily masked through the estimation of a linear model. In the forecasting chapter, running a linear regression model on a clearly nonlinear data set still resulted in statistically valid models and provided decent estimates. Another specification error that is fairly common has to do with autocorrelated and seasonal data sets. Estimating the demand of flowers in a floral chain without accounting for the holidays (Valentine's Day, Mother's Day, and so forth) is a blatant specification error. Failure to clearly use the correct model specification or first sanitizing the data may result in highly erroneous results.

4. Omitted and Redundant Variables This type of model error in multivariate regression exists when regression is used to forecast the future. Suppose an analyst uses multivariate regression to obtain a statistical relationship between a dependent variable (e.g., sales, prices, revenues) and other regressors or independent variables (e.g., economic conditions, advertising levels, market competition) and he or she hopes to use this relationship to forecast the future. Unfortunately, the analyst may not have all the available information at his or her fingertips. If important information is unavailable, an important variable may be omitted (e.g., market saturation effects, price elasticity of demand, threats of emerging technology), or if too much data is available, redundant variables may be included in the analysis (e.g., inflation rate, interest rate, economic growth). It may be counterintuitive, but the problem of redundant variables is more serious than omitted variables.[1] In a situation where redundant variables exist,[2] and if these redundant variables are perfectly correlated or collinear with each other, the regression equation does not exist and cannot be solved. In the case where slightly less severe collinearity exists, the estimated regression equation will be less accurate than without this collinearity. For instance, suppose both interest rates and inflation rates are used as explanatory variables in the regression analysis, where if there is a significant negative correlation between these variables with a time lag, then using both variables to explain sales revenues in the future is redundant. Only one variable is sufficient to explain the relationship with sales. If the analyst uses both variables, the errors in the regression analysis will increase. The prediction errors of an additional variable will increase the errors of the entire regression.

5. Heteroskedasticity If the variance of the errors in a regression analysis increases over time, the regression equation is said to be flawed and suffers from heteroskedasticity. Although this may seem to be a technical matter, many regression practitioners fall into this heteroskedastic trap without even realizing it. See Chapter 9, Using the Past to Predict the Future, for details on heteroskedasticity, testing for its existence, and methods to fix the error.

6. Multicollinearity One of the assumptions required for a regression to run is that the independent variables are noncorrelated or noncollinear. These independent variables are exactly collinear when a variable is an exact linear combination of the other variables. This error is most frequently encountered when dummy variables are used.[3] A quick check of multicollinearity is to run a correlation matrix of the independent variables.[4] In most instances, the multicollinearity problem will prevent the regression results from being computed. See Chapter 9's Appendix D—Detecting and Fixing Multicollinearity—for more details.

7. Spurious Regression, Data Mining, Time Dependency, and Survivorship Bias Spurious regression is another danger that analysts often run into. This mistake is made through certain uses of data-mining activities. Data mining refers to using approaches such as a step-wise regression analysis, where analysts do not have some prior knowledge of the economic effects of what independent variables drive the dependent variable, and use all available data at their disposal. The analyst then runs a step-wise regression, where the methodology ranks the highest correlated variable to the least correlated variable.[5] Then the methodology automatically adds each successive independent variable in accordance with its correlation until some specified stopping statistical criteria. The resulting regression equation is then taken as the final and best result. The problem with this approach is that some independent variables may simply be randomly moving about while the dependent variable may also be randomly moving about, and their movements depend on time.[6] Suppose this randomness in motion is somehow related at certain points in time but the actual economic fundamentals or financial relationships do not exist. Data-mining activities will pick up the coincidental randomness and not the actual relationship, and the result is a spurious regression; that is, the relationship estimated is bogus and is purely a chance happenstance. Multicollinearity effects may also unnecessarily eliminate highly significant variables from the step-wise regression.

Finally, survivorship bias and self-selection bias are important, as only the best-performing realization will always show up and have the most amount of visibility. For instance, looking to the market to obtain proxy data can be dangerous for only successful firms will be around and have the data. Firms that have failed will most probably leave no trails of their existence, let alone credible market data for an analyst to collect. Self-selection occurs when the data that exist are biased and selective. For instance, pharmacology research on a new cancer treatment will attract cancer patients of all types, but the researchers will clearly only select those patients in the earlier stages of cancer, making the results look more promising than they actually are.

8. Autoregressive Processes, Lags, Seasonality, and Serial Correlation In time-series data, certain variables are autoregressive in nature; that is, future values of variables such as price, demand, interest rates, inflation rates, and so forth depend on values that occurred in the past, or are autoregressive.[7] This reversion to the past occurs because of many reasons, including seasonality and cyclicality.[8] Because of these cyclical or seasonal and autoregressive effects, regression analysis using seasonal or cyclical independent variables as is will yield inexact results. In fact, some of these autoregressive, cyclical, or seasonal variables will affect the dependent variable differently over time. There may be a time lag between effects. For example, an increase in interest rates may take 1 to 3 months before the mortgage market feels the effects of this decline. Ignoring this time lag will downplay the relationships of highly significant variables.

9. Correlation and Causality Regression analysis looks at correlation effects, not causality.[9] To say that there is a cause in *X* (independent variable) that drives the outcome of *Y* (dependent variable) through the use of regression analysis is flawed. For instance, there is a high correlation between the number of shark attacks and lunch hour around the world. Clearly, sharks cannot tell that it is time to have lunch. However, because lunchtime is the warmest time of the day, this is also the hour that beaches around the world are most densely populated. With a higher population of swimmers, the chances of heightened shark attacks are almost predictable. Lunchtime does not *cause* sharks to go hungry and prompt them to search for food. Just because there is a correlation does not mean that there is causality. Making this leap will provide analysts and management an incorrect interpretation of the results.

10. Random Walks Certain financial data (e.g., stock prices, interest rates, inflation rates) follow something called a random walk. Random walks can take on different characteristics, including random walks with certain jumps, random walks with a drift rate, or a random walk that centers or reverts to some long-term average value. Even the models used to estimate random walks are varied, from geometric to exponential, among other things. A simple regression equation will yield no appreciable relationship when random walks exist.[10]

11. Jump Processes Jump processes are more difficult to grasp but are nonetheless important for management to understand and challenge the assumptions of an analyst's results. For instance, the price of oil in the global market may sometimes follow a jump process. When the United States goes to war with another country, or when OPEC decides to cut the production of oil by several billion barrels a year, oil prices will see a sudden jump.

Forecasting revenues based on these oil prices over time using historical data may not be the best approach. These sudden probabilistic jumps should most certainly be accounted for in the analysis. In this case, a jump-diffusion stochastic model is more appropriate than simple time-series or regression analyses.

12. Stochastic Processes Other stochastic processes are also important when analyzing and forecasting the future. Interest rates and inflation rates may follow a mean-reversion stochastic process; that is, interest rates and inflation rates cannot increase or decrease so violently that they fall beyond all economic rationale. In fact, economic factors and pressures will drive these rates to their long-run averages over time. Failure to account for these effects over the long run may yield statistically incorrect estimates, resulting in erroneous forecasts.

> Warning signs to watch out for in time-series forecasting and regression as well as questions to ask include whether the forecasts are out of range, are there structural and business breaks anticipated in the forecast period, are there any misspecifications in the model, are there any possibilities of omitted and redundant variables, are there heteroskedasticity effects, are there any spurious relationships and biases, are there autoregressive lags, are correlations confused with causality, and are there variables that follow a random walk, jump processes, or other stochastic processes.

READING THE WARNING SIGNS IN REAL OPTIONS ANALYSIS

Risk analysis is never complete without the analysis of real options. What are uncertainty and risk analyses good for if one cannot make use of them? Real options analysis looks at the flexibility of a project or management's ability to make midcourse corrections when uncertainty becomes resolved over time. At the outset, real options analysis looks like a very powerful analytical tool, but care should be taken when real options analysis is applied. For instance, consider the following.

1. Do Not Let Real Options Simply Overinflate the Value of a Project One of the most significant criticisms of real options approaches is that of overinflating the value of a project. This criticism, of course, is false. Real options are applicable if and only if the following requirements are met: traditional financial analysis can be performed and models can be built; uncertainty exists; the same uncertainty drives value; management or the project has strategic options or flexibility to either take advantage of these uncertainties or to hedge them; and management must be credible in executing the relevant strategic options when they become optimal to do so; otherwise all the options in the world would be useless. Thus, an analyst should not simply apply real options analysis to every project that comes across his or her desk, but only to those that are appropriate and ripe for analysis.

An option will always bear a value greater than or equal to zero. Hence, critics argue that by applying real options analysis, a project's value will be artificially inflated. In reality, real options may sometimes appear without cost, but in most cases firms need to pay to acquire these options (e.g., spending money to retrofit a refinery to obtain a switching option to choose between input fuels), and although the value of an option may be positive, its value can be clouded by the cost to obtain the option, making the entire strategy unprofitable and reducing the value of a project. So, although the value of an option is positive, the entire strategy's value may be negative. The lesson here is well learned—do not apply real options analysis on everything in sight, just to those projects that actually do have strategic options. Without doing so may mean leaving money on the table.

2. How Is Volatility Obtained and How Do You Reconcile Its Value? Fifty percent of the value of a real options analysis is simply thinking about it and realizing that management has the flexibility to make midcourse corrections when uncertainty becomes resolved over time. Twenty-five percent is crunching the numbers and the remaining 25 percent of the value in applying real options comes from being able to convince and explain the results to management. One of the toughest things to explain is the concept of where and how volatility is obtained. Volatility should be obtained from a project based on a project's level of uncertainty going forward. One major error is to use external market proxies for volatility. Using a firm's stock price to estimate volatility of a single project in a company with hundreds or even thousands of projects is not only incorrect, it is ludicrous. An analyst should, hence, be able to defend the choice of volatility estimates. See Johnathan Mun's *Real Options Analysis, Second Edition* (Wiley, 2005) for details on converting volatility to probability, and explaining volatility to management in an easy to understand manner.

3. What About Competing Options or Options That Have Not Even Been Considered? If a project has 10 strategic options, do you analyze all 10 options? What about projects in the distant future, where the options are not yet known for certain, that may be highly valuable? For a project with many options, the analyst has to determine which of these options are independent and which are interacting type options. If the options are interacting, dominant strategies will always dominate over less valuable options and the value of the project's total set of options will revert to these dominant options.[11] Thus, do not evaluate all the options in the world if only a few options capture a significant portion of the value. Focus instead on valuing those important or dominant options.

4. The Error of Interpretation of Option Results Sometimes options come without a cost, while sometimes they do have a cost. On some occasions, option value is tangible or explicit, and sometimes option value is implicit or intangible. As an example, the land seller illustration used in Chapter 13, The Black Box Made Transparent: Real Options Super Lattice Solver, looks at the value of having an abandonment option, where if the counterparty signs the contractual agreement, the maximum expected cost of the contract is the option value.[12] However, in the case of some of the illustrations in Chapter 12, What's So Real About Real Options, and Why Are They Optional? where a research and development outfit performing stage-gate development has the option to abandon at every stage, valuing these options does not automatically mean the IRS or a counterparty will show up at the door and give the company a check in that amount. In this situation, the option value is an intangible or implicit value, useful as a measure against other projects and alternate strategies with or without such a flexibility option value.[13]

> Warning signs to watch out for in real options analysis and questions to ask include whether the real options analysis is applied inappropriately when there are no options such that the value of a project is inappropriately overinflated, how the volatility measure is obtained, are competing or omitted options appropriately considered, and are the results interpreted correctly.

READING THE WARNING SIGNS IN OPTIMIZATION UNDER UNCERTAINTY

Finally, uncertainty and risk analyses are irrelevant if these quantified risks cannot be diversified away. Optimization looks at the ability to diversify away risks to find the best combination of projects subject to some prespecified constraints.

1. Why Are the Decision Variables the Decision Variables? Decision variables are the variables that management has control over (e.g., which projects to execute, which products to manufacture, which vendor to purchase from, which wells to drill). However, sometimes things that are seemingly decision variables on the outset may not exactly be decision variables. For instance, the CEO's pet project is definitely a "go" decision no matter what the analytical results. The internal politics involved in decision making is something that cannot be taken lightly. Decision variables in an optimization analysis should most certainly be decision variables, not decisions that have already been made with the façade that their existence still has to be justified. Finally, certain decision variables are related to other decision variables and this interaction must be considered. For instance, Project A is a precursor to Projects B, C, and D; however, Project C cannot be executed if project B is executed, and Project C is a precursor to Project D.[14]

2. How Certain Are the Optimization Results? Has the analyst looked at enough combinations to obtain the optimal results? In static optimization without simulation, whether it is using Risk Simulator, Excel's goal seek, or Excel's Solver add-in, the optimal solution will be found, if there is one, rather quickly, as the computer can calculate all possible combinations and permutations of inputs to yield the optimal results. However, in optimization under uncertainty,[15] the process will take much longer and the results may not achieve optimality quickly. Even if the results do seem to be optimal, it is hard to tell; thus, it is safer to run the simulation much longer than required. An impatient analyst may fall into the trap of not running sufficient simulation trials to obtain robust stochastic or dynamic optimization results.

3. What Is the Analyst's Level of Training? Little knowledge of probability will lead to more dangerous conclusions than no knowledge at all. Knowledge and experience together will prove to be an impressive combination, especially when dealing with advanced analytics. Almost always, the first step in getting more advanced analytics accepted and rolled out corporate-wide is to have a group of in-house experts trained in both the art and science of advanced analytics. Without a solid foundation, plans on rolling out these analytics will fail miserably.

> Warning signs to watch out for in an optimization under uncertainty and questions to ask include whether the decision variables are indeed decisions to be made, what are the levels of certainty of the results, and what is the level of training of the analyst.

QUESTIONS

1. Define what is meant by negligent entrustment.
2. What are some of the general types of errors encountered by an analyst when creating a model?
3. Why is truncation in a model's assumption important? What would happen to the results if truncation is not applied when it should be?
4. What is a critical success factor?
5. What are some of the normal-looking statistics?
6. What are structural breaks and specification errors, and why are they important?

CHAPTER 16

Changing a Corporate Culture

HOW TO GET RISK ANALYSIS ACCEPTED IN AN ORGANIZATION

Advanced analytics is hard to explain to management.[1] So, how do you get risk analysis accepted as the norm in a corporation, especially if your industry is highly conservative? It is almost a guarantee in conservative companies that an analyst showing senior management a series of fancy, mathematically complex, and computationally sophisticated models will be thrown out of the office together with his or her results and have the door slammed in his or her face. Changing management's thinking is the topic of discussion in this chapter. Explaining results and convincing management appropriately go hand in hand with the characteristics of the advanced analytical tools, which if they satisfy certain change-management requisites, the level and chances of acceptance become easier.

CHANGE-MANAGEMENT ISSUES AND PARADIGM SHIFTS

Change-management specialists have found that there are several criteria to be met before a paradigm shift in thinking is found to be acceptable in a corporation. For example, in order for senior management to accept a new and novel set of advanced analytical approaches—simulation, forecasting, real options, and portfolio optimization—the models and processes themselves must have applicability to the problem at hand, and not merely be an academic exercise.[2] Figure 16.1 lists the criteria required for change.

As we saw previously, it is certainly true that large multinationals have embraced the concept of risk analysis with significant fervor, and that risk analysis is here to stay.[3] It is not simply an academic exercise, nor is it the

"No change of paradigm comes easily"

Criteria for instituting change:

- Method applicability
  - Not just an academic exercise
- Accurate, consistent, and replicable
  - Creates a standard for decision making
- Value-added propositions
  - Competitive advantage over competitors
  - Provide valuable insights otherwise unavailable
- Exposition
  - Making the black box transparent
  - Explaining the value to senior management
- Comparative advantage
  - Better method than the old
  - It takes a good theory to kill an old one
- Compatibility with the old approach
  - Based on the old with significant improvements
- Flexibility
  - Able to be tweaked
  - Covers a multitude of problems
- External influences
  - From "Main Street" to "Wall Street"
  - Communicating to the investment community the value created internally

FIGURE 16.1 Changing a corporate culture.

latest financial analysis fad that is here today and gone tomorrow. In addition, the process and methodology have to be consistent, accurate, and replicable, that is, they pass the scientific process. Given similar assumptions, historical data, and assertions, one can replicate the results with ease and predictability. This replicability is especially true with the use of software programs such as the ones included on the CD-ROM.

Next, the new method must provide a compelling value-added proposition. Otherwise, it is nothing but a fruitless and time-consuming exercise. The time, resources, and effort spent must be met and even surpassed by the method's added value. This added value is certainly the case in larger capital investment initiatives, where a firm's future or the future of a business unit may be at stake—incorrect and insufficient results may be obtained, and disastrous decisions made if risk analysis is not undertaken.

Other major criteria include the ability to provide the user a comparative advantage over its competitors, which is certainly the case when the additional valuable insights generated through advanced risk analysis will help management identify options, value, prioritize, and select strategic and less risky alternatives that may otherwise be overlooked.

Finally, in order to accept a change in mind-set, the new methodology, analysis, process, or model must be easy to explain and understand. In addition, there has to be a link to previously accepted methods, whether the new methodology is an extension of the old or a replacement of the old due to some clear superior attributes. These last two points are the most difficult to tackle for an analyst. The sets of criteria prior to this are direct and easy to define.

The new set of risk analytics is nothing but an extension of existing methodologies.[4] For instance, Monte Carlo simulation can be explained simply as scenario analysis applied to the *n*th degree. Simulation is nothing but scenario analysis done thousands of times only not just on a single variable (e.g., the three common scenarios: good economy, average economy, and bad economy complete with their associated probabilities of occurrence and payoffs at each state), but on multiple variables interacting simultaneously, where multiple variables are changing independently or dependently, in a correlated or uncorrelated fashion (e.g., competition, economy, market share, technological efficacy, and so forth). In fact, the result stemming from new analytics is simply a logical extension of the traditional approaches. Figure 16.2 illustrates this logical extension.

The static model in the illustration shows a revenue value of $2, cost of $1, and the resulting income value, calculated as the difference between the two, of $1. Compare that to the dynamic model, where the same inputs are used but the revenue and cost variables have been subjected to Monte Carlo simulation. Once simulation has been completed, the dynamic model still shows the same single-point estimate of $1 as in the static model. In other words, adding in the more advanced analytics, namely, Monte Carlo simulation, the model and results have not changed. If management still wants the single-point estimate of $1 reported, then so be it. However, by logical extension, if both revenues and costs are uncertain, then by definition, the resulting income will also be uncertain. The forecast chart for the income variable shows this uncertainty of the resulting income with fluctuations around $1. In fact, additional valuable information is obtained using simulation, where the probability or certainty of breakeven or exceeding $0 in income is shown as 95.40 percent in Figure 16.2. In addition, rather than relying on the single-point estimate of $1, simulation reveals that the business only has an 8.90 percent probability of exceeding the single-point estimate of $1 in income (Figure 16.3).

STATIC MODEL			DYNAMIC MODEL			
Revenue	$	2.00	Revenue	$	2.00	<<--- This is an Input Assumption
Cost	$	1.00	Cost	$	1.00	<<--- This is an Input Assumption
Income	$	1.00	Income	$	1.00	<<--- This is an Output Forecast

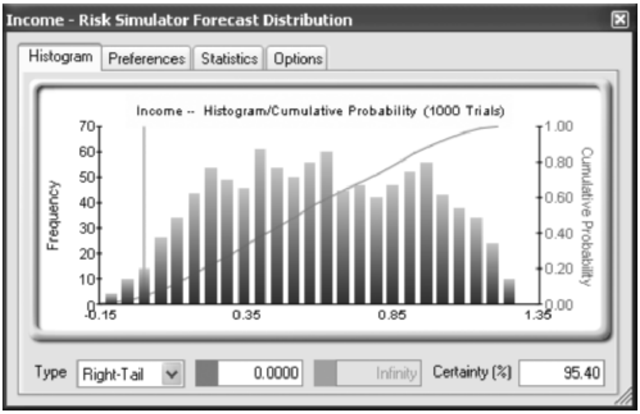


FIGURE 16.2 Monte Carlo simulation as a logical extension of traditional analysis.

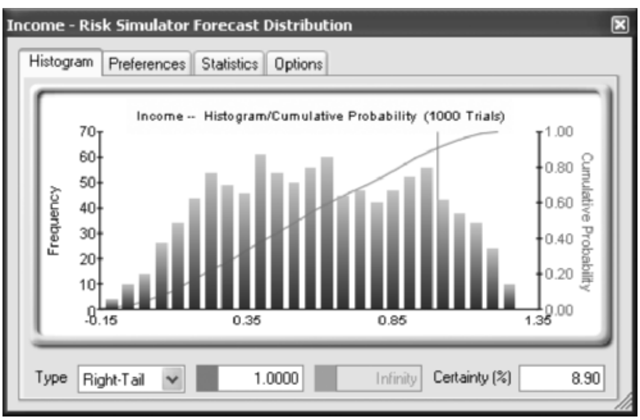


FIGURE 16.3 Probability of exceeding the original $1 income.

If simulation is not applied here, the riskiness of this project will never be clearly elucidated. Imagine if management has multiple but similar types of projects where every project has a single-point estimate of $1. In theory, management should be indifferent in choosing any of these projects. However, if the added element of risk is analyzed, each project may have different probabilities of breakeven and different probabilities of exceeding the $1 income threshold. Clearly, the project with the least amount of risk should be chosen (i.e., highest probability of breakeven and exceeding the threshold value).

MAKING TOMORROW'S FORECAST TODAY

Firms that are at first skeptical about applying advanced analytics in their decision-making activities should always consider first applying these new rules to smaller projects. Instead of biting off too much immediately, a small-scale project is always preferable. Companies new to advanced risk analytics should first learn to crawl before they start running and head straight for the wall. If management can be eased into the new analytical paradigm slowly, the transition will be more palatable.

Having a vision to change the entire organization's decision-making processes overnight is very admirable but will be very short-lived and bound for disaster. Before an organization can learn to make tomorrow's forecast today, it has to learn from the lessons of yesterday. One approach is to look at high-profile projects in the past. Instead of starting with forecasting, perform some back-casting first. Instead of waiting for years to verify if the results from the analysis were actually correct or valuable, the result from a back-casting analysis is almost immediate. If the analyst is true to himself or herself, using the actual data coupled with the assumptions used in the past (without the advantage of hindsight), the new analytical results can then be compared to the decisions that were made to see if different strategies and decisions would have been undertaken instead. However, care should be taken as corporate politics come into play because the individuals who made the decisions in the past may not take it too kindly when their decisions are negatively scrutinized.

No matter the strategy moving forward, one thing is certain: If senior management buys into the techniques, acceptance will be imminent. Otherwise, a few junior analysts in a cubicle somewhere trying to get management's attention will fail miserably. In retrospect, a midlevel manager trying to impress his or her superiors without the adequate knowledge and support from analysts will not work either.

The approach for successful implementation has to be comprehensive and three pronged. Senior management must keep an open mind to alternatives. Middle management must keep championing the approach and not let minor setbacks be permanent, while attempting to be the conduit of information between the junior analysts and senior management. Finally, analysts should attempt to acquire as much knowledge about the techniques and applications as possible. The worst possible outcome is where extreme expectations are set from high above and the powers that be, while the lower rungs cannot deliver the goods as required.

In order to facilitate adoption of a new set of analytical methods in an organization, several criteria must first be met. To judge the level of potential adoption, the following factors should be considered: whether the method is applicable to the problem at hand, how accurate, consistent, and replicable are the methods, what are the value-added propositions, what is the level of expositional ease, what are the comparative advantages, how compatible is the new method to the old models, how flexible is the new method, and what are some of the external influences in using the methods.

Notes

Chapter 1 Moving Beyond Uncertainty

1. Peter L. Bernstein, *Against the Gods: The Remarkable Story of Risk* (John Wiley & Sons, 1996).
2. Save the potentiality of a plane crash, at which I would have very much regretted not taking the parachute.
3. The concepts of high risk and high return are nothing new and are central to the development of the *capital asset pricing model (CAPM)* used to estimate the required rate of return on a project based on its *systematic risk*. In the CAPM model, the higher the risk, the higher the expected rate of return (ceteris paribus, or holding everything else constant).
4. Risk can be measured in different ways. In this example, it is measured using the standard deviation of the distribution of returns.
5. This selection is because Project X bears a positive net return (positive net present value) above its implementation cost, making it profitable. Thus, the cheapest project is selected.
6. "Independence" means that the projects themselves are uncorrelated; thus it is assumed that there are no risk-diversification effects. "Mutually exclusive" means that the manager cannot mix and match among the different projects (e.g., 2 Project Xs with 3 Project Ys).
7. This choice of course is based purely on financial analysis alone by holding everything else constant (management's taste and preferences, or other strategic values inherent in different projects).
8. On a continuous basis, the probability of hitting exactly \$30 (i.e., \$30.0000000000 and so forth) is very close to zero. The probability in a distribution is measured as the area under the curve, which means two values are required, for example, the probability of net revenues being between \$29 and \$31 is 25 percent. Thus, the area under the curve for a single-point estimate (a single line in a distribution) is close to zero.
9. The *Law of Demand* in economics requires that, in most cases, price and quantity demanded are negatively correlated, in accordance with a downward-sloping demand curve. The exception being Giffen or status goods where a higher price may yield a higher quantity demanded (e.g., Porsches are desirable and have a higher status because they are expensive, among other things).

10. A firm's average variable cost curve is U shaped, with an initial downward slope at lower quantities (economies of scale), hits a global minimum value where marginal cost equals average variable cost, and then continues to slope upward (diseconomies of scale).
11. See Chapter 9 for details on time-series and regression models.
12. The simulated actual values depicted graphically are based on a Geometric Brownian Motion with a volatility of 20 percent calculated as the standard deviation of the simulated natural logarithms of historical returns.
13. See Chapters 2 and 3 for details of other measures of risk and uncertainty.

Chapter 2 From Risk to Riches

1. Ron Dembo and Andrew Freeman, *Seeing Tomorrow: Rewriting the Rules of Risk* (John Wiley & Sons, 1998). This book provides an interesting nonmathematical review of risk management.
2. That is, the standard deviation of the population (σ) and the standard deviation of a sample (s) are

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}} \qquad s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

   where the standard deviation is the square root of the sum of the deviation of each data point (x_i) from the population mean (μ) or sample mean ($\bar{x}$) squared, and then divided into the population size (N) or sample size (n) less one. For the sample statistic, the division is into n less one to correct for the degrees of freedom in a smaller sample size. The variance is simply the square of the standard deviation.
3. Johnathan Mun, *Real Options Analysis, Second Edition* (John Wiley & Sons, 2005); Johnathan Mun, *Real Options Analysis Course* (John Wiley & Sons, 2003). Refer to these books for details on estimating volatility in a real options context.
4. For instance, the height distribution's mean is 10 m with a standard deviation of 1 m, which yields a coefficient variation of 0.1, versus the weight distribution's mean of 100 kg and a standard deviation of 20 kg, which yields a coefficient of variation of 0.2. Clearly, the weight distribution carries with it more variability.

Chapter 3 A Guide to Model-Building Etiquette

1. However, be aware that password busters are abundant and certain spreadsheet models can very easily be hacked by outsiders. A better

approach is to convert sensitive functions and macros into ActiveX "*.dll*" files that are encrypted, providing a much higher level of security.

Chapter 4 On the Shores of Monaco

1. This example is an adaptation from papers and lectures provided by Professor Sam Savage of Stanford University.
2. In this example, the median is a better measure of central tendency.
3. The same nonparametric simulation can also be applied using Risk Simulator's custom distribution where each occurrence has an equal chance of being selected.
4. The approach used here is the application of a Geometric Brownian Motion stochastic process for forecasting and simulating potential outcomes.

Chapter 5 Test Driving Risk Simulator

1. This approach is valid because in typical simulations, thousands of trials are being simulated, and the assumption of normality can be applied.

Chapter 7 Extended Business Cases I: Pharmaceutical and Biotech Negotiations, Oil and Gas Exploration, Financial Planning with Simulation, Hospital Risk Management, and Risk-Based Executive Compensation Valuation

1. For example, drilling engineers can review historical drilling cost data and provide a probability distribution of drilling costs in a geographic area in the proposed rock formation. They are not required to know how important this risk is to the project economics versus the risk that an oil and gas reservoir is not present after the well is drilled. This risk is better evaluated by geological/geophysical staff.
2. While "low risk" is a subjective term, the risk in our model reflects a well that might be drilled in or very close to an existing producing oil field in a mature, well-established oil basin such as the Permian Basin of West Texas.
3. The economic limit is the point at which the marginal expense of producing the well exceeds the marginal revenue associated with the oil or gas produced. It is highly dependent on the company's organization and producing infrastructure. For our model we assume 10 BOPD is the economic limit.
4. Calculated from average weekly prices of West Texas Intermediate Crude, then averaged over 52 weeks of each year, from November 1991 to March 2003.

5. Note that if it is determined that the well has not encountered significant oil and gas reserves, the well is not completed and these costs are not incurred. This cost is the only one of the Year 0 costs in our model that is not incurred in the case of a dry hole.
6. NPV/I is simply the net present value of the project divided by the sum of Year 0 investments, and provides a measure of bang for the buck in a capital-rationing corporate environment.
7. In fact, most oil and gas companies do maintain proprietary price forecasts for the purpose of portfolio and investment analysis. Sensitivity of projects to these forecasts suggests that corporations (not just project teams) are well advised to model the variability in earnings and cash flow that will propagate from unavoidable errors in their proprietary price forecasts.

Chapter 9 Using the Past to Predict the Future

1. An arbitrary 3-month moving average is chosen. For modeling purposes, different *n*-length moving averages should be computed and the one with the least amount of errors should be chosen.
2. To start Excel's Data Analysis, first click on the *Tools* menu in Excel and select *Add-Ins*. Then make sure the check box beside *Analysis Tool Pak* is selected and hit *OK*. Then return to the *Tools* menu and select *Data Analysis*. The *Regression* functionality should now exist.
3. See Chapter 8, Making Tomorrow's Forecast Today, for specifics on using Risk Simulator.
4. The critical *t-statistic* can be found in the t-distribution table at the end of this book, by looking down the two-tailed alpha 0.025 (alpha 0.05 for two tails means that each tail has an area of 0.025) and cross-referencing it to 6 degrees of freedom. The degrees of freedom is calculated as the number of data points, *n*, (7) used in the regression, less the number of independent regressors, *k* (1).
5. As this is a two-tailed hypothesis test, the alpha should be halved, which means that as long as the p-value calculated is less than 0.025 (half of 0.05), then the null hypothesis should be rejected.
6. The adjusted R-squared is used here as this is a multivariate regression, and the adjustment in the coefficient of determination accounts for the added independent variable.
7. The two most notable and challenging econometric models include the ARCH (autoregressive conditional heteroskedasticity) and GARCH (generalized autoregressive conditional heteroskedasticity) models.

Chapter 10 The Search for the Optimal Decision

1. For a total cost of $550 for the entire trip.

2. The number of possible itineraries is the factorial of the number of cities, that is, $3! = 3 \times 2 \times 1 = 6$.
3. A total of five cities means $5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$.
4. A triangular distribution can be applied here, with the minimum level set at \$300, most likely a value of \$325, and a maximum level set at \$500.
5. The straight lines in Figure 10.6 would now be nonlinear and the problem would be difficult to solve graphically.
6. To access Solver, start Excel, click on *Tools | Add-Ins*. Make sure the check box beside *Solver Add-In* is selected. Solver can then be accessed by clicking on *Tools | Solver*.
7. A two-decision variable optimization problem requires a two-dimensional graph, which means an *n*-decision variable problem requires the use of an *n*-dimensional graph, making the problem mathematically and manually intractable using the graphical method.
8. Chapter 11, Optimization under Uncertainty, illustrates a similar portfolio-optimization process but under uncertainty using Risk Simulator.

Chapter 14 Extended Business Cases II: Real Estate, Banking, Military Strategy, Automotive Aftermarkets, Global Earth Observation Systems, and Employee Stock Options

1. McKinsey & Company, Inc., Tom Copeland, Tim Koller, and Jack Murrin, *Valuation: Measuring and Managing the Value of Companies,* 3rd ed. (John Wiley & Sons, 2000).
2. United States. President. The National Security Strategy of the United States of America. Washington, White House, 2002. Internet: 1 June, 2005: http://www.whitehouse.gov/nsc/nss.pdf p. 30.
3. United States. President. The National Security Strategy of the United States of America. Washington, White House, 2002. Internet: 1 June, 2005: http://www.whitehouse.gov/nsc/nss.pdf p. 30.
4. Chizek, Judy G., "Military Transformation: Intelligence, Surveillance and Reconnaissance," Congressional Research Services, Jan. 17, 2003, p. 2.
5. Housel, T., O. El Sawy, J. Zhong, and W. Rodgers. "Models for Measuring the Return on Information Technology: A Proof of Concept Demonstration." 22nd International Conference on Information Systems. December, 2001. p. 13.
6. Ibid.
7. The ISR Mission is generally conducted at a highly classified level, so specifics of the ICP and CCOP are not available to the public. For the purpose of this academic research, much of the data was estimated or inferred based on realistic sampling of unclassified process information.

Information on human capital, such as salaries and operator training, are public information and were gathered from sources such as the Stay Navy Website and the Center for Information Dominance (CID) training documentation. The equipment data was also derived or inferred from documentation provided by the OPNAV N20 staff and the Space and Naval Warfare Command (SPAWAR). Other information such as number of process outputs and executions was extrapolated from samples gathered via interviews with ISR crews currently operating on board deployed U.S. Navy surface ships.

8. *Wall Street Journal*, April 21, 2004.
9. Financial Accounting Standards Board web site: www.fasb.org.
10. See Johnathan Mun's *Real Options Analysis*, Second Edition (Wiley Finance, 2005) for details on the case study.
11. The GBM accounts for dividends on European options, but the basic BSM does not.
12. American options are exercisable at any time up to and including the expiration date. European options are exercisable only at termination or maturity expiration date. Most ESOs are a mixture of both—European option during the vesting period (the option cannot be exercised prior to vesting) reverting to an American option after the vesting period.
13. These could be cliff vesting (the options are all void if the employee leaves or is terminated before this cliff vesting period) or graded monthly/quarterly/annually vesting (a certain proportion of the options vest after a specified period of employment service to the firm).
14. The BSM described herein refers to the original model developed by Fisher Black, Myron Scholes, and Robert Merton. Although significant advances have been made such that the BSM can be modified to take into consideration some of the exotic issues discussed in this case study, it is mathematically very complex and is highly impractical for use.
15. This multiple is the ratio of the stock price when the option is exercised to the contractual strike price, and is tabulated based on historical information. Post- and near-termination exercise behaviors are excluded.
16. For instance, a 1,000-step nonrecombining binomial lattice will require 2×10^{301} computations, and even after combining all of the world's fastest supercomputers together, will take longer than the lifetime of the sun to compute!
17. A tornado chart lists all the inputs that drive the model, starting from the input variable that has the most effect on the results. The chart is obtained by perturbing each input at some consistent range (e.g., ±10 percent from the base case) one at a time, and comparing their results to the base case. Different input levels yield different tornado charts, but in most cases, volatility is not the only dominant variable. Forfeiture, vesting, and suboptimal exercise behavior multiples all tend to either dominate over or be as dominant as volatility.

18. A spider chart looks like a spider with a central body and its many legs protruding. The positively sloped lines indicate a positive relationship (e.g., the higher the stock price, the higher the option value), while a negatively sloped line indicates a negative relationship. Further, spider charts can be used to visualize linear and nonlinear relationships.
19. People tend to exhibit suboptimal exercise behavior due to many reasons, for example, the need for liquidity, risk adversity, personal preferences, and expectations.
20. Of the 6,553 stocks analyzed, 2,924 of them pay dividends, with 2,140 of them yielding at or below 5 percent, 2,282 at or below 6 percent, 2,503 at or below 7 percent, and 2,830 at or below 10 percent.
21. An alternative method is to calculate the relevant carrying cost adjustment by artificially inserting an inflated dividend yield to convert the ESO into a "soft option," thereby discounting the value of the ESO. This method is more difficult to apply and is susceptible to more subjectivity than using a put option.
22. Cedric Jolidon finds the mean values of marketability discounts to be between 20 percent and 35 percent in his article, "The Application of the Marketability Discount in the Valuation of Swiss Companies" (Swiss Private Equity Corporate Finance Association). A typical marketability range of 10–40% was found in several discount court cases. In the *CPA Journal* (February 2001), M. Greene and D. Schnapp found that a typical range was somewhere between 30% and 35%. Another article in the *Business Valuation Review* finds that 35 percent is the typical value (Jay Abrams, "Discount for Lack of Marketability"). In the *Fair Value* newsletter, Michael Paschall finds that 30–50% is the typical marketability discount used in the market.
23. Any level of precision and confidence can be chosen. Here, the 99.9 percent statistical confidence with a $0.01 error precision ($0.01 fluctuation around the average option value) is fairly restrictive. Of course, the level of precision attained is contingent on the inputs and their distributional parameters being accurate.
24. This assumes that the inputs are valid and accurate.
25. A 1,000-step customized binomial lattice is generally used unless otherwise noted. Sometimes increments from 1,000 to 5,000 steps may be used to check for convergence. However, due to the nonrecombining nature of changing volatility options, a lower number of steps may have to be employed.
26. This proprietary algorithm was developed by Dr. Johnathan Mun based on his analytical work with FASB in 2003–2004; his books: *Valuing Employee Stock Options Under the 2004 FAS 123 Requirements* (Wiley Finance, 2004), *Real Options Analysis: Tools and Techniques, Second Edition* (Wiley Finance, 2005), *Real Options Analysis Course* (Wiley Finance, 2003), and *Applied Risk Analysis: Moving Beyond Uncertainty*

(Wiley Finance, 2003); creation of his software, Real Options Super Lattice Solver; academic research; and previous valuation consulting experience at KPMG Consulting.

27. A nonrecombining binomial lattice bifurcates (splits into two) every step it takes, so starting from one value, it branches out to two values on the first step (2^1), two becomes four in the second step (2^2), and four becomes eight in the third step (2^3), and so forth, until the 1,000th step (2^{1000} or over 10^{301} values to calculate; the world's fastest supercomputer cannot calculate the result within our lifetimes).
28. The Law of Large Numbers stipulates that the central tendency (mean) of a distribution of averages is an unbiased estimator of the true population average. The results from 4,200 steps show a mean value that is comparable to the median of the distribution of averages, and, hence, 4,200 as the number of steps is chosen as the input into the binomial lattice.

Chapter 15 The Warning Signs

1. The problem of omitted variables is less vital as an analyst will simply have to work with all available data. If everything about the future is known, then why bother forecasting? If there is no uncertainty, then the future is known with certainty.
2. The problem of redundant variables is also known as multicollinearity.
3. For instance, if a dummy variable on sex is used (i.e., "0" for male and "1" for female), then a regression equation with *both* dummy variables will be perfectly collinear. In such a situation, simply drop one of the dummy variables as they are mutually exclusive of each other.
4. Make sure there are no independent variables that are perfectly or almost perfectly correlated to each other. In addition, correlation analysis can be performed to test the linear relationships among the independent variables.
5. If Y is the dependent variable and X_i is the independent variable, then the correlation pairs are between all possible combinations of Y and X_i.
6. Interest rates tend to be time dependent (mean-reverting over longer periods of time) and demand for a product that is not related to interest rate movements may also be time dependent (exhibiting cyclicality and seasonality effects).
7. The term "auto" means self and "regressive" means reverting to the past. Hence, the term "autoregressive" means to revert back to one's own past history.
8. Seasonality effects are usually because of periodicities in time (12-month seasonality in a year, 4-quarter seasonality in a year, 7-day seasonality in a week, etc.) while cyclical effects are because of larger influences

without regard to periodicities (e.g., business cycle movements and technological innovation cycles).

9. However, there are other approaches used to estimate causality, for example, Granger causality approaches look at statistical causalities.
10. More advanced econometric models are required to estimate random walks, including methods using differences and unit root models.
11. See *Real Options Analysis, Second Edition* (Wiley, 2005) and *Real Options Analysis Course* (Wiley, 2003) for details on interacting options in evaluating the chooser option.
12. In this case, the option value is explicit, or something that is tangible, and the seller of the option can actually acquire this value.
13. Therefore, management's compensation should not be tied to actualizing this implicit option value.
14. This means the decision strategies are: A–B, A–C, and A–C–D.
15. Optimization under uncertainty means to run a set of simulations for a certain number of trials (e.g., 1,000 trials), pause, estimate the forecast distributions, test a set of combinations of decision variables, and rerun the entire analysis again, for hundreds to thousands of times.

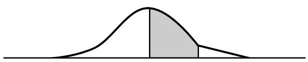
Chapter 16 Changing a Corporate Culture

1. Advanced analytics are all the applications discussed in this book, including simulation, time-series forecasting, regression, optimization, and real options analysis.
2. Examples of an academic exercise that has little pragmatic application for general consumption in the areas of advanced analytics include sensitivity simulation, variance reduction, closed-form partial-differential models, and so forth. These are mathematically elegant approaches, but they require analysts with advanced degrees in finance and mathematics to apply, making the methodology and results very difficult to explain to management.
3. The case is made through the many actual business cases and examples throughout this book.
4. This is particularly true for Monte Carlo simulation where simulation cannot be applied unless there already is a spreadsheet model.

Tables You Really Need

Standard Normal Distribution (partial area)
Standard Normal Distribution (full area)
Student's t-Distribution (one tail and two tails)
Durbin–Watson Critical Values (alpha 0.05)
Normal Random Numbers (standard normal distribution's random number generated ~ ($N(0,1)$)
Random Numbers (multiple digits)
Uniform Random Numbers (uniform distribution's random number generated between 0.0000 and 1.0000)
Chi-Square Critical Values
F-Distribution Critical Statistics (alpha one tail 0.10)
F-Distribution Critical Statistics (alpha one tail 0.05)
F-Distribution Critical Statistics (alpha one tail 0.25)
F-Distribution Critical Statistics (alpha one tail 0.01)
Real Options Analysis Values (1-year maturity at 5% risk-free rate)
Real Options Analysis Values (3-year maturity at 5% risk-free rate)
Real Options Analysis Values (5-year maturity at 5% risk-free rate)
Real Options Analysis Values (7-year maturity at 5% risk-free rate)
Real Options Analysis Values (10-year maturity at 5% risk-free rate)
Real Options Analysis Values (15-year maturity at 5% risk-free rate)
Real Options Analysis Values (30-year maturity at 5% risk-free rate)

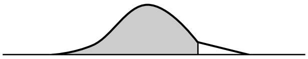
Standard Normal Distribution (partial area)



Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990

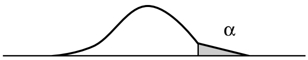
Example: For a Z-value of 1.96, refer to the 1.9 row and 0.06 column for the area of 0.4750. This means there is 47.50% in the shaded region and 2.50% in the single tail. Similarly, there is 95% in the body or 5% in both tails.

Standard Normal Distribution (full area)



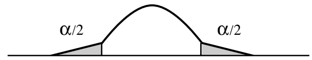
Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

Example: For a Z-value of 2.33, refer to the 2.3 row and 0.03 column for the area of 0.99. This means there is 99% in the shaded region and 1% in the one-sided left or right tail.



one tail

Student's t-Distribution (one and two tails)



two tail

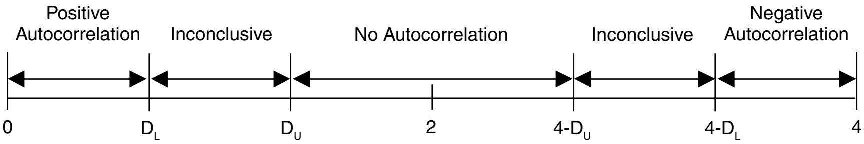
alpha	0.1	0.05	0.025	0.01	0.005
df = 1	3.0777	6.3137	12.7062	31.8210	63.6559
2	1.8856	2.9200	4.3027	6.9645	9.9250
3	1.6377	2.3534	3.1824	4.5407	5.8408
4	1.5332	2.1318	2.7765	3.7469	4.6041
5	1.4759	2.0150	2.5706	3.3649	4.0321
6	1.4398	1.9432	2.4469	3.1427	3.7074
7	1.4149	1.8946	2.3646	2.9979	3.4995
8	1.3968	1.8595	2.3060	2.8965	3.3554
9	1.3830	1.8331	2.2622	2.8214	3.2498
10	1.3722	1.8125	2.2281	2.7638	3.1693
15	1.3406	1.7531	2.1315	2.6025	2.9467
20	1.3253	1.7247	2.0860	2.5280	2.8453
25	1.3163	1.7081	2.0595	2.4851	2.7874
30	1.3104	1.6973	2.0423	2.4573	2.7500
35	1.3062	1.6896	2.0301	2.4377	2.7238
40	1.3031	1.6839	2.0211	2.4233	2.7045
45	1.3007	1.6794	2.0141	2.4121	2.6896
50	1.2987	1.6759	2.0086	2.4033	2.6778
100	1.2901	1.6602	1.9840	2.3642	2.6259
200	1.2858	1.6525	1.9719	2.3451	2.6006
300	1.2844	1.6499	1.9679	2.3388	2.5923
500	1.2832	1.6479	1.9647	2.3338	2.5857
100000	1.2816	1.6449	1.9600	2.3264	2.5759

alpha	0.1	0.05	0.025	0.01	0.005
df = 1	6.3137	12.7062	25.4519	63.6559	127.3211
2	2.9200	4.3027	6.2054	9.9250	14.0892
3	2.3534	3.1824	4.1765	5.8408	7.4532
4	2.1318	2.7765	3.4954	4.6041	5.5975
5	2.0150	2.5706	3.1634	4.0321	4.7733
6	1.9432	2.4469	2.9687	3.7074	4.3168
7	1.8946	2.3646	2.8412	3.4995	4.0294
8	1.8595	2.3060	2.7515	3.3554	3.8325
9	1.8331	2.2622	2.6850	3.2498	3.6896
10	1.8125	2.2281	2.6338	3.1693	3.5814
15	1.7531	2.1315	2.4899	2.9467	3.2860
20	1.7247	2.0860	2.4231	2.8453	3.1534
25	1.7081	2.0595	2.3846	2.7874	3.0782
30	1.6973	2.0423	2.3596	2.7500	3.0298
35	1.6896	2.0301	2.3420	2.7238	2.9961
40	1.6839	2.0211	2.3289	2.7045	2.9712
45	1.6794	2.0141	2.3189	2.6896	2.9521
50	1.6759	2.0086	2.3109	2.6778	2.9370
100	1.6602	1.9840	2.2757	2.6259	2.8707
200	1.6525	1.9719	2.2584	2.6006	2.8385
300	1.6499	1.9679	2.2527	2.5923	2.8279
500	1.6479	1.9647	2.2482	2.5857	2.8195
100000	1.6449	1.9600	2.2414	2.5759	2.8071

Example: For an alpha in the single right tail area of 2.5% with 15 degrees of freedom, the critical t value is 2.1315.

Durbin–Watson Critical Values (alpha 0.05)

	k = 1		k = 2		k = 3		k = 4		k = 5	
n	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	1.67	0.90	1.83	0.79	1.99
21	1.22	1.42	1.13	1.54	1.03	1.66	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	1.65	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



Example: For 30 observations (n) of a multivariate regression with three independent variables, the critical Durbin–Watson statistics are 1.21 (D_L) and 1.65 (D_U). If the calculated Durbin–Watson is 1.05, there is positive autocorrelation.

Normal Random Numbers (standard normal distribution's random number generated ~ $N(0,1)$)

	1	2	3	4	5	6	7	8	9	10
1	-1.0800	-0.5263	-0.7099	-0.3124	0.0216	-0.7768	-0.0752	0.4273	0.7708	0.1887
2	-1.1028	1.0904	-0.9228	-0.8881	-1.7909	0.6459	0.8982	-0.9736	-0.8630	0.1361
3	-0.8336	0.1454	-1.5907	1.0843	0.6271	1.1925	1.4669	0.5701	-2.7364	0.2500
4	0.2296	-0.2436	-0.0639	0.2307	-0.0560	-1.8494	0.6068	-0.2562	0.2168	-0.0261
5	1.2795	-0.6267	0.3133	0.3831	0.8894	0.9869	1.6185	0.7713	0.1421	-0.9623
6	1.2079	-0.8924	0.0491	0.0250	-0.5501	-0.8312	0.5067	-0.4316	0.7880	0.3858
7	-0.9474	-1.1758	-2.0242	-1.1567	-0.3838	0.8031	-0.5129	1.3572	-0.6772	1.0510
8	-0.7296	-0.8073	0.1137	-0.3553	-2.5826	-0.2768	0.0300	0.6233	-2.0171	-1.0818
9	0.0939	-0.1833	0.5550	0.3809	0.4096	0.0930	0.0257	-0.0603	-2.3620	-0.2656
10	-1.2110	-0.3240	0.8859	0.3776	-1.9103	2.0585	0.5215	-1.3543	-0.6975	-1.5965
11	-0.4614	-0.7827	0.8294	0.4460	-0.6563	0.4167	-0.3699	-0.0852	0.5010	0.3579
12	-0.5282	1.2526	-0.3289	1.5912	0.8460	1.2919	-0.6255	-0.2466	0.6740	1.6007
13	1.1204	0.5921	0.3115	0.1986	-0.6793	0.0694	-0.2777	0.5517	-0.5385	1.2437
14	-0.3726	0.0955	-2.3786	-1.7042	0.6656	0.0641	0.3874	1.1669	-0.6837	-0.0934
15	-0.5656	-0.0949	-0.3845	-0.6864	0.9967	0.0695	1.4614	1.0945	-1.2097	-1.4070
16	-0.2430	-2.4107	-2.5924	0.2724	-0.0967	-0.0315	-0.8218	0.2390	0.5987	-0.6879
17	-0.2820	-0.4370	0.7358	-0.3511	-0.2308	-0.7651	-0.7652	-0.4937	-1.0157	-0.1394
18	-0.3955	0.5096	0.1447	-0.4119	1.3781	-0.7365	0.4475	1.7877	0.3629	1.4260
19	0.1652	-0.4687	0.1058	-0.4183	-0.3782	-2.4017	0.9160	-1.8322	-0.6279	0.0098
20	-0.0504	-1.0931	-1.6450	-0.6165	-0.0279	-0.9539	-1.6489	-0.7252	0.3962	0.8928
21	0.1841	-0.1236	0.7653	-0.9054	0.8158	-0.8576	1.9970	-0.1568	-1.6658	-0.6698
22	-1.1091	0.5140	0.4505	-1.7429	0.0854	0.1573	-2.2687	0.4879	-0.0820	0.4840
23	0.6553	0.4692	0.9139	0.9639	-0.9046	-0.6695	-0.3393	-1.8453	1.0532	0.9795
24	0.5185	0.8624	0.6098	0.7062	0.3533	0.1695	0.1840	-0.5235	0.7202	0.0790
25	-0.6228	-0.0052	0.1012	0.9541	1.4046	-0.2620	-0.2783	0.7601	-0.0375	1.8253
26	0.5867	0.3346	-0.0588	-0.4356	0.0004	0.2037	-1.1411	-0.4674	2.2770	-0.8338

27	0.2450	1.0948	-0.8954	1.0444	-0.2184	-1.1320	1.5127	-0.9275	-0.4799	0.1281
28	-0.0279	-0.1937	-1.2914	-0.9880	1.1571	0.5578	0.4071	1.2601	1.1695	-0.2957
29	-0.4161	-0.5507	-0.4475	0.0689	0.4422	-1.1679	-0.5163	0.3915	-0.7226	0.9784
30	-0.8053	0.3502	-1.4505	-0.5941	-0.7228	-0.7034	-1.0992	0.3020	-0.1026	-1.2502
31	1.0404	0.1097	0.4544	-0.5799	-0.2926	1.2725	-0.5619	-0.0821	-0.5477	1.0231
32	0.2528	0.5059	-1.4190	0.3989	-1.3937	-1.2064	0.0228	-0.6627	1.1379	0.5220
33	-0.2739	-0.9455	-2.2941	0.0276	1.7592	-1.7925	-0.5070	-0.2650	1.5300	-0.3373
34	-0.9423	0.3491	-1.3512	0.4576	1.0860	-0.1653	0.4558	-0.6405	-1.2085	-0.7493
35	0.0883	0.2888	-0.5136	2.1450	-0.0262	2.9286	-1.7310	1.1511	-0.6439	-0.3583
36	-0.4517	0.2437	0.2776	-0.7868	0.1671	1.0155	-0.3549	0.7456	-0.3971	-1.9802
37	-1.1278	-2.3892	-0.2134	0.2925	1.2178	-0.3160	0.9686	-1.2743	-0.0707	1.5162
38	1.3791	-0.4170	-0.1155	-0.1992	-1.1890	1.2458	-1.6882	0.3428	-1.3231	-0.3701
39	0.0819	0.5604	-1.7606	-0.6743	-1.0426	-0.8501	1.1497	0.0442	0.5657	-1.2778
40	-0.4175	0.4203	1.2675	1.2768	-0.4826	-2.3268	0.0747	1.0223	0.2681	-0.3952
41	0.6801	-0.6346	-0.4628	0.1047	1.0032	-1.4099	0.3401	-0.5051	-1.2245	-0.4696
42	0.9200	-0.4411	1.9065	-0.8623	-0.8896	-1.3154	-0.2427	1.4517	0.6037	0.7206
43	-2.0794	-0.0927	1.0023	-0.2296	-0.6263	-0.7918	-0.6372	2.7211	0.3840	-0.5358
44	0.5448	0.6405	0.3647	-1.9654	-1.8430	-0.4946	-0.6691	1.3191	0.9991	1.6156
45	1.0963	1.2051	0.7243	2.3032	-0.4820	2.0831	0.6108	0.8796	0.5527	0.8128
46	-0.9386	1.2509	-2.1745	-0.4204	-0.6400	-1.0716	0.0190	-1.9153	-1.4322	0.0870
47	2.4524	1.5695	-0.6953	-2.4997	-0.0891	-0.5719	-0.9301	-0.3394	-2.6532	-0.0226
48	0.4448	-1.8947	0.7942	0.3552	-0.4288	1.0699	0.7316	-1.1951	1.4356	0.2318
49	0.1323	-0.0470	1.5664	0.1610	0.4068	-1.1848	-1.2338	0.1546	-0.3490	2.4516
50	-0.6323	1.7106	-0.6715	0.2511	0.7708	-0.6902	0.8453	1.1715	1.4897	0.0401

Random Numbers (multiple digits)

	1	2	3	4	5	6	7	8	9	10
1	2721.5177	7927.3605	5509.2000	7755.4229	8910.1600	9583.6638	9063.9590	8043.2820	9974.8278	7685.4216
2	5427.5197	6573.0674	6996.6637	9135.8127	2718.8760	8982.9624	4576.1065	5844.0620	2435.0249	2281.9131
3	1570.2192	5024.3217	6764.9039	1023.2814	6548.8675	3329.6628	4520.5547	9269.9768	6344.4565	2809.3591
4	6617.8598	5903.1769	7002.3606	2085.2144	4792.4796	6844.4960	8697.2448	6543.3337	2982.5475	6500.9816
5	1042.3463	4784.7013	2453.3249	2006.1324	2128.2118	4070.3922	7223.9221	9040.6234	3864.3067	8258.1458
6	5152.2026	2683.5095	3648.9192	7937.8332	9361.6421	6588.8570	9066.6720	7688.1069	4799.6166	4936.4821
7	8092.3323	5697.0313	7446.0071	3138.0076	3274.5303	3064.3907	9283.3996	3169.3531	5119.0202	9799.3380
8	3200.8200	8155.2797	2903.4796	3975.4799	2090.4880	2584.4027	2321.1790	9201.1671	5563.8958	4922.1343
9	2182.8695	9863.5501	3827.3677	5479.6807	5846.3606	7009.6787	1956.4793	9485.3016	6048.6349	4545.7721
10	4929.3461	1009.5500	6692.1558	6563.8505	6478.1138	1457.2554	2607.2569	1772.3479	1130.7805	9296.8716
11	9478.2765	9055.7916	8831.3015	9113.9356	3863.2465	6845.3370	7956.4931	3620.3660	6516.1395	5908.0984
12	1336.6521	6161.7270	8222.4781	5859.3163	8247.4744	8348.0894	6487.8202	6784.5221	4693.3882	5667.7078
13	7460.0083	5643.3684	2422.6688	6932.7146	2091.5401	3917.1395	5129.1433	1218.7031	8785.2712	7050.3969
14	5849.9114	5882.0649	6661.0100	6681.4560	9481.2436	2195.1850	4813.7851	9085.3021	1653.4790	3719.3843
15	6975.6430	9691.5555	6668.4537	7785.4196	6508.2217	9147.5266	9760.7188	1920.8204	1278.8593	5578.9917
16	9178.9897	3759.3978	3947.4711	6015.4509	2645.6605	9933.6472	8250.7021	4046.9983	2472.1532	6918.8681
17	1105.8190	7150.7795	1707.7886	6093.6588	5725.3097	6168.8648	6322.9949	8035.1053	1670.8308	3130.7888
18	4378.4322	8484.7097	7236.2981	4585.2984	6117.0657	1604.2704	6441.3144	9050.4318	1192.4602	3053.1196
19	6589.7603	8938.1669	5639.4775	9210.1063	3355.4245	5526.0291	2033.4076	8997.8637	6921.5642	9584.1109
20	4455.2372	6786.2862	4018.8972	5491.1575	1560.0462	4115.5836	1048.3373	9623.0486	8862.3072	7621.1737
21	4448.7636	6209.5568	9959.7063	3177.2467	1641.8797	6802.2869	8161.5705	1685.3721	1941.6971	8308.9046
22	8654.4590	9343.3206	6653.9854	9692.1930	3929.0176	4784.0031	4596.3431	6587.9375	9035.8024	1517.4567
23	8844.1890	9681.9999	2822.7265	2899.0180	5158.0016	5636.0479	2528.0603	6982.0078	9200.3319	6361.4182
24	7390.3963	5983.7082	5900.1055	3837.7891	8828.4116	7731.9270	3157.1180	1957.9680	6105.4342	4370.7669
25	5203.6897	4338.6493	4776.3189	1129.9635	1273.6261	8183.6248	2281.0786	2374.5525	2381.9855	3381.8613
26	4161.9959	6863.8237	9514.2372	2225.4123	4676.0563	6451.0761	5920.1725	2916.4971	5819.8761	7904.7086

27	8044.9759	8610.4069	8708.4209	8303.2069	6696.6600	5799.8857	9579.7723	6845.5490	1039.0858	8763.7395
28	2587.5116	5853.4249	4388.2114	5526.0319	2061.3728	4644.3832	1388.8595	5890.9486	3907.8750	4141.8542
29	7052.2487	6036.5176	2541.3818	2812.2029	7546.7513	2546.8478	2494.0563	6029.0624	1324.0261	8162.4338
30	1163.9374	1931.4068	6247.8204	7745.6642	3070.3767	5071.9130	6159.3637	3013.1682	1226.8873	8162.4898
31	1714.6545	1523.8375	8509.5616	8306.2575	9657.2873	3120.0271	6688.3472	5159.6344	3671.7474	7133.5930
32	3919.0191	5588.6388	4923.3729	5347.2862	1600.1555	2029.1451	3136.0774	2317.8933	3932.5034	3018.1371
33	5026.6414	2547.0444	4424.7295	4170.3210	7624.0027	5232.2546	5874.4753	4124.2614	5273.4984	5929.6120
34	5621.6736	5358.4125	2870.7415	6454.0855	8476.0039	2736.0572	6719.2599	2753.2847	4911.0976	1791.5700
35	9910.5203	6121.8213	3308.4460	3150.5253	9211.2410	5499.6467	3931.8208	5313.8206	5934.1154	4849.0388
36	4312.4452	8426.2265	8872.6974	1663.1930	9120.5661	5981.5407	2613.1288	5439.7424	9611.6777	5188.4457
37	7626.7677	5387.7439	2935.0787	8309.7795	8246.8356	2074.3136	5736.0131	3286.1149	8836.6044	7193.7667
38	9884.8894	6400.4452	3674.4606	6779.5470	9832.8283	8108.7365	4803.5534	7599.8840	2362.3725	8762.0338
39	3835.9843	9103.7538	2867.9787	6320.5689	2208.7881	3409.7276	7836.5953	5104.6250	3424.2561	6521.1725
40	1337.4881	5372.2827	4089.9067	9875.2185	1422.7835	6058.8479	5847.4650	8856.1609	1258.4403	3044.5864
41	3785.1585	4943.2358	2420.7229	5821.4256	5122.6017	8973.0601	6324.9297	9036.6863	1197.9443	2913.1478
42	8926.2929	6024.7767	4233.9264	2292.9495	1958.8263	5534.9091	8243.9129	9370.1993	9628.7974	2321.1484
43	8900.4838	4553.4951	8777.2026	8809.0081	6170.5223	4601.8483	6653.8133	8002.2370	2871.4168	4085.8857
44	6385.5759	1642.6064	4939.2942	8710.5348	2064.4489	7854.8362	4247.8259	3799.9352	6065.6772	9917.2978
45	9684.0981	5429.2985	6042.3134	5461.5755	8034.5336	5056.7885	1621.9722	9290.8556	4395.0623	9808.8263
46	6385.0021	5007.4273	2845.5347	1898.6996	9031.1549	9874.3671	2061.4315	5221.1304	4624.0654	8847.7553
47	8706.7893	1279.1794	8722.8166	5683.5057	9611.4135	2593.5565	2220.1057	6559.8872	3554.2664	5352.1678
48	4542.1757	5609.2758	9599.0981	7644.5129	8663.3854	7009.6717	1887.5296	7330.2408	9197.2417	3012.4571
49	3242.9899	8305.9299	6439.2860	7130.1905	6503.2924	5736.9502	3489.9470	3671.3190	2925.8024	7207.2956
50	7751.6580	7934.9861	8400.8779	2923.6741	4305.5792	3995.4573	9288.3303	6593.9721	5302.2203	9007.2129

Uniform Random Numbers (uniform distribution's random number generated between 0.0000 and 1.0000)

	1	2	3	4	5	6	7	8	9	10
1	0.8470	0.8006	0.8185	0.5479	0.6664	0.4772	0.8983	0.9434	0.0272	0.1912
2	0.8538	0.1840	0.0235	0.5733	0.5103	0.9165	0.2052	0.6861	0.4069	0.8930
3	0.4816	0.0929	0.0404	0.1688	0.4297	0.1381	0.5717	0.3440	0.3050	0.3347
4	0.1827	0.6090	0.2067	0.0201	0.1809	0.4326	0.5870	0.4826	0.8274	0.4693
5	0.6736	0.7903	0.0910	0.7829	0.9657	0.3531	0.5095	0.4019	0.9799	0.4321
6	0.9953	0.8069	0.5096	0.8088	0.5747	0.5876	0.6151	0.7627	0.3793	0.4698
7	0.7613	0.8829	0.9609	0.6287	0.0849	0.9027	0.2761	0.5469	0.5634	0.0308
8	0.1317	0.7907	0.5440	0.0469	0.7220	0.5695	0.2482	0.3742	0.1409	0.3288
9	0.5269	0.6977	0.4061	0.0950	0.2114	0.4113	0.7619	0.6854	0.1402	0.2956
10	0.9121	0.5435	0.3236	0.6256	0.7646	0.3120	0.8037	0.1198	0.8887	0.5443
11	0.5390	0.4622	0.3459	0.1427	0.7762	0.8186	0.5059	0.1905	0.8696	0.8893
12	0.9055	0.4771	0.6290	0.8068	0.5124	0.9142	0.6397	0.5279	0.2051	0.1220
13	0.6644	0.9212	0.2139	0.3678	0.8107	0.1869	0.5594	0.8278	0.2343	0.9175
14	0.7403	0.1068	0.9122	0.1193	0.5645	0.9703	0.9102	0.3528	0.6891	0.0330
15	0.8611	0.9607	0.1820	0.8349	0.4017	0.2822	0.3624	0.8583	0.1495	0.1532
16	0.4914	0.1137	0.2635	0.6062	0.1728	0.5471	0.1065	0.4250	0.7094	0.3168
17	0.7664	0.6767	0.5264	0.9354	0.9880	0.1942	0.9594	0.2610	0.9933	0.3406
18	0.0126	0.5592	0.3942	0.4020	0.7840	0.8675	0.1734	0.0476	0.3372	0.4067
19	0.5251	0.8027	0.6730	0.9985	0.4706	0.2960	0.3305	0.1006	0.1012	0.4638
20	0.7772	0.4434	0.1596	0.3856	0.0163	0.5783	0.4055	0.1490	0.7172	0.2243
21	0.8973	0.7618	0.4225	0.9524	0.7371	0.3863	0.2146	0.3799	0.8521	0.7857
22	0.1709	0.1966	0.1125	0.1454	0.0325	0.2262	0.3624	0.3600	0.6517	0.4073
23	0.1785	0.6833	0.9630	0.3603	0.8863	0.4362	0.5985	0.2979	0.6837	0.0957
24	0.5644	0.2031	0.9500	0.0418	0.9262	0.6584	0.5958	0.9879	0.4332	0.0198
25	0.3672	0.4599	0.2637	0.9380	0.8343	0.6933	0.4732	0.5802	0.2715	0.1287
26	0.8391	0.1803	0.4345	0.7670	0.5298	0.7905	0.4120	0.9688	0.8540	0.8267

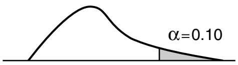
27	0.7135	0.8772	0.5661	0.4345	0.8710	0.6183	0.1704	0.3377	0.1432	0.9205
28	0.9477	0.0880	0.0476	0.2050	0.5699	0.5680	0.3438	0.9242	0.1429	0.0283
29	0.2862	0.0944	0.0698	0.6541	0.5945	0.5464	0.1861	0.8030	0.8177	0.8099
30	0.9237	0.5355	0.9374	0.4701	0.8763	0.3914	0.5917	0.6042	0.0596	0.2829
31	0.5876	0.2458	0.6085	0.6830	0.5682	0.9463	0.5392	0.0854	0.7900	0.3149
32	0.0677	0.4571	0.6932	0.0656	0.3131	0.9006	0.8570	0.7966	0.4101	0.5311
33	0.9369	0.3878	0.8473	0.9510	0.9292	0.1164	0.4611	0.7247	0.7077	0.0106
34	0.1777	0.1686	0.1624	0.9553	0.2083	0.9768	0.2229	0.1562	0.6361	0.0027
35	0.4455	0.5007	0.0395	0.4937	0.9753	0.3447	0.0391	0.6322	0.3977	0.4147
36	0.4002	0.5214	0.1770	0.8398	0.2889	0.5151	0.4960	0.6892	0.4331	0.8813
37	0.4288	0.7095	0.6115	0.1138	0.7932	0.7117	0.6252	0.1275	0.6600	0.0738
38	0.3327	0.3886	0.6723	0.0747	0.7562	0.2142	0.1860	0.9814	0.0407	0.7521
39	0.5113	0.4232	0.2029	0.9034	0.0154	0.6591	0.0515	0.8867	0.5985	0.0338
40	0.2530	0.2622	0.2013	0.0351	0.1554	0.4416	0.0300	0.7017	0.4546	0.6329
41	0.3086	0.7557	0.6003	0.5604	0.6615	0.8889	0.2757	0.8436	0.1147	0.2306
42	0.7732	0.6118	0.3301	0.7272	0.4494	0.4960	0.6787	0.2748	0.4064	0.1111
43	0.6713	0.2170	0.5049	0.7975	0.6739	0.9117	0.0948	0.9233	0.6709	0.6739
44	0.9708	0.0705	0.0987	0.5948	0.1022	0.1206	0.2131	0.3548	0.0826	0.7013
45	0.4756	0.6014	0.8200	0.5208	0.3044	0.4410	0.1012	0.5467	0.7132	0.2751
46	0.6130	0.0888	0.2238	0.1298	0.5416	0.7280	0.9447	0.6551	0.0112	0.5960
47	0.2792	0.7500	0.3124	0.0277	0.3785	0.9622	0.7501	0.6412	0.1556	0.1384
48	0.5724	0.0308	0.7103	0.1949	0.9440	0.9585	0.4508	0.3737	0.7383	0.6845
49	0.2825	0.9384	0.6804	0.3165	0.1243	0.6089	0.2623	0.8008	0.2408	0.9563
50	0.3294	0.4181	0.5703	0.4162	0.8578	0.3346	0.5491	0.1812	0.7001	0.6394

Chi-Square Critical Values

df	0.10	0.09	0.08	0.07	0.06	0.05	0.04	0.03	0.02	0.01
1	2.7055	2.8744	3.0649	3.2830	3.5374	3.8415	4.2179	4.7093	5.4119	6.6349
2	4.6052	4.8159	5.0515	5.3185	5.6268	5.9915	6.4377	7.0131	7.8241	9.2104
3	6.2514	6.4915	6.7587	7.0603	7.4069	7.8147	8.3112	8.9473	9.8374	11.3449
4	7.7794	8.0434	8.3365	8.6664	9.0444	9.4877	10.0255	10.7119	11.6678	13.2767
5	9.2363	9.5211	9.8366	10.1910	10.5962	11.0705	11.6443	12.3746	13.3882	15.0863
6	10.6446	10.9479	11.2835	11.6599	12.0896	12.5916	13.1978	13.9676	15.0332	16.8119
7	12.0170	12.3372	12.6912	13.0877	13.5397	14.0671	14.7030	15.5091	16.6224	18.4753
8	13.3616	13.6975	14.0684	14.4836	14.9563	15.5073	16.1708	17.0105	18.1682	20.0902
9	14.6837	15.0342	15.4211	15.8537	16.3459	16.9190	17.6083	18.4796	19.6790	21.6660
10	15.9872	16.3516	16.7535	17.2026	17.7131	18.3070	19.0208	19.9219	21.1608	23.2093
11	17.2750	17.6526	18.0687	18.5334	19.0614	19.6752	20.4120	21.3416	22.6179	24.7250
12	18.5493	18.9395	19.3692	19.8488	20.3934	21.0261	21.7851	22.7418	24.0539	26.2170
13	19.8119	20.2140	20.6568	21.1507	21.7113	22.3620	23.1423	24.1249	25.4715	27.6882
14	21.0641	21.4778	21.9331	22.4408	23.0166	23.6848	24.4854	25.4931	26.8727	29.1412
15	22.3071	22.7319	23.1992	23.7202	24.3108	24.9958	25.8161	26.8480	28.2595	30.5780
16	23.5418	23.9774	24.4564	24.9901	25.5950	26.2962	27.1356	28.1908	29.6332	31.9999
17	24.7690	25.2150	25.7053	26.2514	26.8701	27.5871	28.4449	29.5227	30.9950	33.4087
18	25.9894	26.4455	26.9467	27.5049	28.1370	28.8693	29.7450	30.8447	32.3462	34.8052
19	27.2036	27.6695	28.1813	28.7512	29.3964	30.1435	31.0367	32.1577	33.6874	36.1908
20	28.4120	28.8874	29.4097	29.9910	30.6488	31.4104	32.3206	33.4623	35.0196	37.5663
21	29.6151	30.0998	30.6322	31.2246	31.8949	32.6706	33.5972	34.7593	36.3434	38.9322
22	30.8133	31.3071	31.8494	32.4526	33.1350	33.9245	34.8672	36.0491	37.6595	40.2894
23	32.0069	32.5096	33.0616	33.6754	34.3696	35.1725	36.1310	37.3323	38.9683	41.6383
24	33.1962	33.7077	34.2690	34.8932	35.5989	36.4150	37.3891	38.6093	40.2703	42.9798
25	34.3816	34.9015	35.4721	36.1065	36.8235	37.6525	38.6417	39.8804	41.5660	44.3140
26	35.5632	36.0914	36.6711	37.3154	38.0435	38.8851	39.8891	41.1461	42.8558	45.6416
27	36.7412	37.2777	37.8662	38.5202	39.2593	40.1133	41.1318	42.4066	44.1399	46.9628
28	37.9159	38.4604	39.0577	39.7213	40.4710	41.3372	42.3699	43.6622	45.4188	48.2782
29	39.0875	39.6398	40.2456	40.9187	41.6789	42.5569	43.6038	44.9132	46.6926	49.5878
30	40.2560	40.8161	41.4303	42.1126	42.8831	43.7730	44.8335	46.1600	47.9618	50.8922
31	41.4217	41.9895	42.6120	43.3033	44.0840	44.9853	46.0595	47.4024	49.2263	52.1914

32	42.5847	43.1600	43.7906	44.4909	45.2815	46.1942	47.2817	48.6410	50.4867	53.4857
33	43.7452	44.3278	44.9664	45.6755	46.4759	47.3999	48.5005	49.8759	51.7429	54.7754
34	44.9032	45.4930	46.1395	46.8573	47.6674	48.6024	49.7159	51.1073	52.9953	56.0609
35	46.0588	46.6558	47.3101	48.0364	48.8560	49.8018	50.9281	52.3350	54.2439	57.3420
36	47.2122	47.8163	48.4782	49.2129	50.0420	50.9985	52.1372	53.5596	55.4889	58.6192
37	48.3634	48.9744	49.6440	50.3869	51.2253	52.1923	53.3435	54.7811	56.7304	59.8926
38	49.5126	50.1305	50.8074	51.5586	52.4060	53.3835	54.5470	55.9995	57.9689	61.1620
39	50.6598	51.2845	51.9688	52.7280	53.5845	54.5722	55.7477	57.2151	59.2040	62.4281
40	51.8050	52.4364	53.1280	53.8952	54.7606	55.7585	56.9459	58.4278	60.4361	63.6908
41	52.9485	53.5865	54.2852	55.0603	55.9345	56.9424	58.1415	59.6379	61.6654	64.9500
42	54.0902	54.7347	55.4405	56.2234	57.1062	58.1240	59.3348	60.8455	62.8918	66.2063
43	55.2302	55.8811	56.5940	57.3845	58.2759	59.3035	60.5257	62.0505	64.1156	67.4593
44	56.3685	57.0258	57.7456	58.5437	59.4436	60.4809	61.7144	63.2531	65.3367	68.7096
45	57.5053	58.1689	58.8955	59.7011	60.6094	61.6562	62.9010	64.4535	66.5552	69.9569
46	58.6405	59.3104	60.0437	60.8568	61.7734	62.8296	64.0855	65.6515	67.7714	71.2015
47	59.7743	60.4503	61.1903	62.0107	62.9355	64.0011	65.2679	66.8475	68.9852	72.4432
48	60.9066	61.5887	62.3353	63.1630	64.0959	65.1708	66.4484	68.0413	70.1967	73.6826
49	62.0375	62.7257	63.4788	64.3137	65.2547	66.3387	67.6270	69.2331	71.4060	74.9194
50	63.1671	63.8612	64.6209	65.4629	66.4117	67.5048	68.8039	70.4229	72.6132	76.1538
51	64.2954	64.9954	65.7615	66.6105	67.5673	68.6693	69.9789	71.6109	73.8183	77.3860
52	65.4224	66.1282	66.9006	67.7567	68.7212	69.8322	71.1521	72.7971	75.0215	78.6156
53	66.5482	67.2598	68.0385	68.9015	69.8737	70.9934	72.3238	73.9813	76.2225	79.8434
54	67.6728	68.3902	69.1751	70.0449	71.0248	72.1532	73.4938	75.1639	77.4217	81.0688
55	68.7962	69.5192	70.3104	71.1870	72.1744	73.3115	74.6622	76.3447	78.6191	82.2920
56	69.9185	70.6472	71.4444	72.3278	73.3227	74.4683	75.8291	77.5239	79.8148	83.5136
57	71.0397	71.7740	72.5773	73.4673	74.4697	75.6237	76.9944	78.7015	81.0085	84.7327
58	72.1598	72.8996	73.7090	74.6055	75.6153	76.7778	78.1583	79.8775	82.2007	85.9501
59	73.2789	74.0242	74.8395	75.7426	76.7597	77.9305	79.3208	81.0520	83.3911	87.1658
60	74.3970	75.1477	75.9689	76.8785	77.9029	79.0820	80.4820	82.2251	84.5799	88.3794

Example: For a degree of freedom (k-c) of 23, the critical values are 32.0069 for 10% alpha level (0.10), 35.1725 for 5% alpha level (0.05), and 41.6383 for 1% alpha level (0.01).

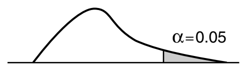


F-Distribution Critical Statistics (alpha one tail 0.10)

	Numerator (df)											
Denominator df	1	2	3	4	5	6	7	8	9	10	15	20
1	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86	60.19	61.22	61.74
2	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.42	9.44
3	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.20	5.18
4	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.87	3.84
5	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30	3.24	3.21
6	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94	2.87	2.84
7	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.70	2.63	2.59
8	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54	2.46	2.42
9	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42	2.34	2.30
10	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.24	2.20
15	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06	1.97	1.92
20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94	1.84	1.79
25	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89	1.87	1.77	1.72
30	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85	1.82	1.72	1.67
35	2.85	2.46	2.25	2.11	2.02	1.95	1.90	1.85	1.82	1.79	1.69	1.63
40	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76	1.66	1.61
45	2.82	2.42	2.21	2.07	1.98	1.91	1.85	1.81	1.77	1.74	1.64	1.58
50	2.81	2.41	2.20	2.06	1.97	1.90	1.84	1.80	1.76	1.73	1.63	1.57
100	2.76	2.36	2.14	2.00	1.91	1.83	1.78	1.73	1.69	1.66	1.56	1.49
200	2.73	2.33	2.11	1.97	1.88	1.80	1.75	1.70	1.66	1.63	1.52	1.46
300	2.72	2.32	2.10	1.96	1.87	1.79	1.74	1.69	1.65	1.62	1.51	1.45
500	2.72	2.31	2.09	1.96	1.86	1.79	1.73	1.68	1.64	1.61	1.50	1.44
100000	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60	1.49	1.42

Example: For an alpha in the single right-tail area of 10% with 10 degrees of freedom in the numerator and 15 degrees of freedom in the denominator, the critical F value is 2.06.

	Numerator (df)										
Denominator df	25	30	35	40	45	50	100	200	300	500	100000
1	62.05	62.26	62.42	62.53	62.62	62.69	63.01	63.17	63.22	63.26	63.33
2	9.45	9.46	9.46	9.47	9.47	9.47	9.48	9.49	9.49	9.49	9.49
3	5.17	5.17	5.16	5.16	5.16	5.15	5.14	5.14	5.14	5.14	5.13
4	3.83	3.82	3.81	3.80	3.80	3.80	3.78	3.77	3.77	3.76	3.76
5	3.19	3.17	3.16	3.16	3.15	3.15	3.13	3.12	3.11	3.11	3.11
6	2.81	2.80	2.79	2.78	2.77	2.77	2.75	2.73	2.73	2.73	2.72
7	2.57	2.56	2.54	2.54	2.53	2.52	2.50	2.48	2.48	2.48	2.47
8	2.40	2.38	2.37	2.36	2.35	2.35	2.32	2.31	2.30	2.30	2.29
9	2.27	2.25	2.24	2.23	2.22	2.22	2.19	2.17	2.17	2.17	2.16
10	2.17	2.16	2.14	2.13	2.12	2.12	2.09	2.07	2.07	2.06	2.06
15	1.89	1.87	1.86	1.85	1.84	1.83	1.79	1.77	1.77	1.76	1.76
20	1.76	1.74	1.72	1.71	1.70	1.69	1.65	1.63	1.62	1.62	1.61
25	1.68	1.66	1.64	1.63	1.62	1.61	1.56	1.54	1.53	1.53	1.52
30	1.63	1.61	1.59	1.57	1.56	1.55	1.51	1.48	1.47	1.47	1.46
35	1.60	1.57	1.55	1.53	1.52	1.51	1.47	1.44	1.43	1.42	1.41
40	1.57	1.54	1.52	1.51	1.49	1.48	1.43	1.41	1.40	1.39	1.38
45	1.55	1.52	1.50	1.48	1.47	1.46	1.41	1.38	1.37	1.36	1.35
50	1.53	1.50	1.48	1.46	1.45	1.44	1.39	1.36	1.35	1.34	1.33
100	1.45	1.42	1.40	1.38	1.37	1.35	1.29	1.26	1.24	1.23	1.21
200	1.41	1.38	1.36	1.34	1.32	1.31	1.24	1.20	1.18	1.17	1.14
300	1.40	1.37	1.34	1.32	1.31	1.29	1.22	1.18	1.16	1.14	1.12
500	1.39	1.36	1.33	1.31	1.30	1.28	1.21	1.16	1.14	1.12	1.09
1000	1.38	1.34	1.32	1.30	1.28	1.26	1.19	1.13	1.11	1.08	1.01

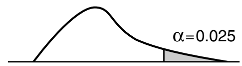


F-Distribution Critical Statistics (alpha one tail 0.05)

	Numerator (df)											
Denominator df	1	2	3	4	5	6	7	8	9	10	15	20
1	161	199	216	225	230	234	237	239	241	242	246	248
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.43	19.45
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.70	8.66
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.86	5.80
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.62	4.56
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	3.94	3.87
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.51	3.44
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.22	3.15
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.01	2.94
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.85	2.77
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.40	2.33
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.20	2.12
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.09	2.01
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.01	1.93
35	4.12	3.27	2.87	2.64	2.49	2.37	2.29	2.22	2.16	2.11	1.96	1.88
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	1.92	1.84
45	4.06	3.20	2.81	2.58	2.42	2.31	2.22	2.15	2.10	2.05	1.89	1.81
50	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07	2.03	1.87	1.78
100	3.94	3.09	2.70	2.46	2.31	2.19	2.10	2.03	1.97	1.93	1.77	1.68
200	3.89	3.04	2.65	2.42	2.26	2.14	2.06	1.98	1.93	1.88	1.72	1.62
300	3.87	3.03	2.63	2.40	2.24	2.13	2.04	1.97	1.91	1.86	1.70	1.61
500	3.86	3.01	2.62	2.39	2.23	2.12	2.03	1.96	1.90	1.85	1.69	1.59
100000	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.67	1.57

	Numerator (df)										
Denominator df	25	30	35	40	45	50	100	200	300	500	100000
1	249	250	251	251	251	252	253	254	254	254	254
2	19.46	19.46	19.47	19.47	19.47	19.48	19.49	19.49	19.49	19.49	19.50
3	8.63	8.62	8.60	8.59	8.59	8.58	8.55	8.54	8.54	8.53	8.53
4	5.77	5.75	5.73	5.72	5.71	5.70	5.66	5.65	5.64	5.64	5.63
5	4.52	4.50	4.48	4.46	4.45	4.44	4.41	4.39	4.38	4.37	4.37
6	3.83	3.81	3.79	3.77	3.76	3.75	3.71	3.69	3.68	3.68	3.67
7	3.40	3.38	3.36	3.34	3.33	3.32	3.27	3.25	3.24	3.24	3.23
8	3.11	3.08	3.06	3.04	3.03	3.02	2.97	2.95	2.94	2.94	2.93
9	2.89	2.86	2.84	2.83	2.81	2.80	2.76	2.73	2.72	2.72	2.71
10	2.73	2.70	2.68	2.66	2.65	2.64	2.59	2.56	2.55	2.55	2.54
15	2.28	2.25	2.22	2.20	2.19	2.18	2.12	2.10	2.09	2.08	2.07
20	2.07	2.04	2.01	1.99	1.98	1.97	1.91	1.88	1.86	1.86	1.84
25	1.96	1.92	1.89	1.87	1.86	1.84	1.78	1.75	1.73	1.73	1.71
30	1.88	1.84	1.81	1.79	1.77	1.76	1.70	1.66	1.65	1.64	1.62
35	1.82	1.79	1.76	1.74	1.72	1.70	1.63	1.60	1.58	1.57	1.56
40	1.78	1.74	1.72	1.69	1.67	1.66	1.59	1.55	1.54	1.53	1.51
45	1.75	1.71	1.68	1.66	1.64	1.63	1.55	1.51	1.50	1.49	1.47
50	1.73	1.69	1.66	1.63	1.61	1.60	1.52	1.48	1.47	1.46	1.44
100	1.62	1.57	1.54	1.52	1.49	1.48	1.39	1.34	1.32	1.31	1.28
200	1.56	1.52	1.48	1.46	1.43	1.41	1.32	1.26	1.24	1.22	1.19
300	1.54	1.50	1.46	1.43	1.41	1.39	1.30	1.23	1.21	1.19	1.15
500	1.53	1.48	1.45	1.42	1.40	1.38	1.28	1.21	1.18	1.16	1.11
100000	1.51	1.46	1.42	1.39	1.37	1.35	1.24	1.17	1.14	1.11	1.01

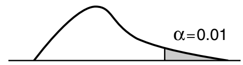
F-Distribution Critical Statistics (alpha one tail 0.025)



	Numerator (df)											
Denominator df	1	2	3	4	5	6	7	8	9	10	15	20
1	648	799	864	900	922	937	948	957	963	969	985	993
2	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40	39.43	39.45
3	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42	14.25	14.17
4	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.66	8.56
5	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.43	6.33
6	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.27	5.17
7	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.57	4.47
8	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.10	4.00
9	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.77	3.67
10	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.52	3.42
15	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06	2.86	2.76
20	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.57	2.46
25	5.69	4.29	3.69	3.35	3.13	2.97	2.85	2.75	2.68	2.61	2.41	2.30
30	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.31	2.20
35	5.48	4.11	3.52	3.18	2.96	2.80	2.68	2.58	2.50	2.44	2.23	2.12
40	5.42	4.05	3.46	3.13	2.90	2.74	2.62	2.53	2.45	2.39	2.18	2.07
45	5.38	4.01	3.42	3.09	2.86	2.70	2.58	2.49	2.41	2.35	2.14	2.03
50	5.34	3.97	3.39	3.05	2.83	2.67	2.55	2.46	2.38	2.32	2.11	1.99
100	5.18	3.83	3.25	2.92	2.70	2.54	2.42	2.32	2.24	2.18	1.97	1.85
200	5.10	3.76	3.18	2.85	2.63	2.47	2.35	2.26	2.18	2.11	1.90	1.78
300	5.07	3.73	3.16	2.83	2.61	2.45	2.33	2.23	2.16	2.09	1.88	1.75
500	5.05	3.72	3.14	2.81	2.59	2.43	2.31	2.22	2.14	2.07	1.86	1.74
100000	5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11	2.05	1.83	1.71

	Numerator (df)										
Denominator df	25	30	35	40	45	50	100	200	300	500	100000
1	998	1001	1004	1006	1007	1008	1013	1016	1017	1017	1018
2	39.46	39.46	39.47	39.47	39.48	39.48	39.49	39.49	39.49	39.50	39.50
3	14.12	14.08	14.06	14.04	14.02	14.01	13.96	13.93	13.92	13.91	13.90
4	8.50	8.46	8.43	8.41	8.39	8.38	8.32	8.29	8.28	8.27	8.26
5	6.27	6.23	6.20	6.18	6.16	6.14	6.08	6.05	6.04	6.03	6.02
6	5.11	5.07	5.04	5.01	4.99	4.98	4.92	4.88	4.87	4.86	4.85
7	4.40	4.36	4.33	4.31	4.29	4.28	4.21	4.18	4.17	4.16	4.14
8	3.94	3.89	3.86	3.84	3.82	3.81	3.74	3.70	3.69	3.68	3.67
9	3.60	3.56	3.53	3.51	3.49	3.47	3.40	3.37	3.36	3.35	3.33
10	3.35	3.31	3.28	3.26	3.24	3.22	3.15	3.12	3.10	3.09	3.08
15	2.69	2.64	2.61	2.59	2.56	2.55	2.47	2.44	2.42	2.41	2.40
20	2.40	2.35	2.31	2.29	2.27	2.25	2.17	2.13	2.11	2.10	2.09
25	2.23	2.18	2.15	2.12	2.10	2.08	2.00	1.95	1.94	1.92	1.91
30	2.12	2.07	2.04	2.01	1.99	1.97	1.88	1.84	1.82	1.81	1.79
35	2.05	2.00	1.96	1.93	1.91	1.89	1.80	1.75	1.74	1.72	1.70
40	1.99	1.94	1.90	1.88	1.85	1.83	1.74	1.69	1.67	1.66	1.64
45	1.95	1.90	1.86	1.83	1.81	1.79	1.69	1.64	1.62	1.61	1.59
50	1.92	1.87	1.83	1.80	1.77	1.75	1.66	1.60	1.58	1.57	1.55
100	1.77	1.71	1.67	1.64	1.61	1.59	1.48	1.42	1.40	1.38	1.35
200	1.70	1.64	1.60	1.56	1.53	1.51	1.39	1.32	1.29	1.27	1.23
300	1.67	1.62	1.57	1.54	1.51	1.48	1.36	1.28	1.25	1.23	1.18
500	1.65	1.60	1.55	1.52	1.49	1.46	1.34	1.25	1.22	1.19	1.14
100000	1.63	1.57	1.52	1.48	1.45	1.43	1.30	1.21	1.17	1.13	1.01

F-Distribution Critical Statistics (alpha one tail 0.01)



Denominator df	Numerator (df) 1	2	3	4	5	6	7	8	9	10	15	20
1	4052	4999	5404	5624	5764	5859	5928	5981	6022	6056	6157	6209
2	98.50	99.00	99.16	99.25	99.30	99.33	99.36	99.38	99.39	99.40	99.43	99.45
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.34	27.23	26.87	26.69
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.20	14.02
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.72	9.55
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.56	7.40
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.31	6.16
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.52	5.36
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	4.96	4.81
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.56	4.41
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.52	3.37
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.09	2.94
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13	2.85	2.70
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.70	2.55
35	7.42	5.27	4.40	3.91	3.59	3.37	3.20	3.07	2.96	2.88	2.60	2.44
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.52	2.37
45	7.23	5.11	4.25	3.77	3.45	3.23	3.07	2.94	2.83	2.74	2.46	2.31
50	7.17	5.06	4.20	3.72	3.41	3.19	3.02	2.89	2.78	2.70	2.42	2.27
100	6.90	4.82	3.98	3.51	3.21	2.99	2.82	2.69	2.59	2.50	2.22	2.07
200	6.76	4.71	3.88	3.41	3.11	2.89	2.73	2.60	2.50	2.41	2.13	1.97
300	6.72	4.68	3.85	3.38	3.08	2.86	2.70	2.57	2.47	2.38	2.10	1.94
500	6.69	4.65	3.82	3.36	3.05	2.84	2.68	2.55	2.44	2.36	2.07	1.92
100000	6.64	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.04	1.88

	Numerator (df)										
Denominator df	25	30	35	40	45	50	100	200	300	500	100000
1	6240	6260	6275	6286	6296	6302	6334	6350	6355	6360	6366
2	99.46	99.47	99.47	99.48	99.48	99.48	99.49	99.49	99.50	99.50	99.50
3	26.58	26.50	26.45	26.41	26.38	26.35	26.24	26.18	26.16	26.15	26.13
4	13.91	13.84	13.79	13.75	13.71	13.69	13.58	13.52	13.50	13.49	13.46
5	9.45	9.38	9.33	9.29	9.26	9.24	9.13	9.08	9.06	9.04	9.02
6	7.30	7.23	7.18	7.14	7.11	7.09	6.99	6.93	6.92	6.90	6.88
7	6.06	5.99	5.94	5.91	5.88	5.86	5.75	5.70	5.68	5.67	5.65
8	5.26	5.20	5.15	5.12	5.09	5.07	4.96	4.91	4.89	4.88	4.86
9	4.71	4.65	4.60	4.57	4.54	4.52	4.41	4.36	4.35	4.33	4.31
10	4.31	4.25	4.20	4.17	4.14	4.12	4.01	3.96	3.94	3.93	3.91
15	3.28	3.21	3.17	3.13	3.10	3.08	2.98	2.92	2.91	2.89	2.87
20	2.84	2.78	2.73	2.69	2.67	2.64	2.54	2.48	2.46	2.44	2.42
25	2.60	2.54	2.49	2.45	2.42	2.40	2.29	2.23	2.21	2.19	2.17
30	2.45	2.39	2.34	2.30	2.27	2.25	2.13	2.07	2.05	2.03	2.01
35	2.35	2.28	2.23	2.19	2.16	2.14	2.02	1.96	1.94	1.92	1.89
40	2.27	2.20	2.15	2.11	2.08	2.06	1.94	1.87	1.85	1.83	1.80
45	2.21	2.14	2.09	2.05	2.02	2.00	1.88	1.81	1.79	1.77	1.74
50	2.17	2.10	2.05	2.01	1.97	1.95	1.82	1.76	1.73	1.71	1.68
100	1.97	1.89	1.84	1.80	1.76	1.74	1.60	1.52	1.49	1.47	1.43
200	1.87	1.79	1.74	1.69	1.66	1.63	1.48	1.39	1.36	1.33	1.28
300	1.84	1.76	1.70	1.66	1.62	1.59	1.44	1.35	1.31	1.28	1.22
500	1.81	1.74	1.68	1.63	1.60	1.57	1.41	1.31	1.27	1.23	1.16
100000	1.77	1.70	1.64	1.59	1.55	1.52	1.36	1.25	1.20	1.15	1.01

Real Options Analysis Values (1-year maturity at 5% risk-free rate)

	Profitability Ratio (% in-the-money)									
Volatility	-99%	-90%	-80%	-70%	-60%	-50%	-40%	-30%	-20%	-10%
1%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
3%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.04%
5%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.35%
7%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.02%	0.88%
9%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.10%	1.52%
11%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%	0.30%	2.22%
13%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.05%	0.60%	2.96%
15%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%	0.13%	1.01%	3.72%
17%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.02%	0.28%	1.49%	4.49%
19%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.06%	0.50%	2.03%	5.26%
21%	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%	0.13%	0.78%	2.63%	6.05%
23%	0.00%	0.00%	0.00%	0.00%	0.00%	0.02%	0.24%	1.13%	3.26%	6.84%
25%	0.00%	0.00%	0.00%	0.00%	0.00%	0.05%	0.40%	1.54%	3.93%	7.63%
27%	0.00%	0.00%	0.00%	0.00%	0.01%	0.11%	0.61%	2.00%	4.62%	8.43%
29%	0.00%	0.00%	0.00%	0.00%	0.02%	0.18%	0.86%	2.50%	5.33%	9.22%
31%	0.00%	0.00%	0.00%	0.00%	0.04%	0.29%	1.17%	3.05%	6.06%	10.02%
33%	0.00%	0.00%	0.00%	0.00%	0.07%	0.44%	1.52%	3.63%	6.80%	10.82%
35%	0.00%	0.00%	0.00%	0.01%	0.12%	0.62%	1.92%	4.24%	7.55%	11.62%
37%	0.00%	0.00%	0.00%	0.02%	0.18%	0.84%	2.35%	4.88%	8.32%	12.42%
39%	0.00%	0.00%	0.00%	0.03%	0.27%	1.10%	2.82%	5.54%	9.09%	13.21%
41%	0.00%	0.00%	0.00%	0.05%	0.39%	1.39%	3.33%	6.22%	9.86%	14.01%
43%	0.00%	0.00%	0.00%	0.08%	0.53%	1.73%	3.87%	6.91%	10.64%	14.80%
45%	0.00%	0.00%	0.01%	0.13%	0.70%	2.10%	4.44%	7.62%	11.43%	15.60%
47%	0.00%	0.00%	0.01%	0.19%	0.91%	2.50%	5.03%	8.35%	12.22%	16.39%
49%	0.00%	0.00%	0.02%	0.26%	1.14%	2.93%	5.64%	9.09%	13.01%	17.18%
51%	0.00%	0.00%	0.03%	0.36%	1.41%	3.40%	6.28%	9.83%	13.80%	17.97%
53%	0.00%	0.00%	0.05%	0.48%	1.71%	3.90%	6.94%	10.59%	14.60%	18.76%
55%	0.00%	0.00%	0.08%	0.62%	2.03%	4.42%	7.61%	11.35%	15.40%	19.55%
57%	0.00%	0.00%	0.11%	0.78%	2.39%	4.96%	8.30%	12.12%	16.20%	20.33%
59%	0.00%	0.00%	0.16%	0.97%	2.78%	5.54%	9.00%	12.90%	16.99%	21.12%
61%	0.00%	0.00%	0.21%	1.18%	3.19%	6.13%	9.72%	13.68%	17.79%	21.90%
63%	0.00%	0.01%	0.29%	1.42%	3.63%	6.74%	10.44%	14.46%	18.59%	22.68%
65%	0.00%	0.01%	0.37%	1.69%	4.10%	7.37%	11.18%	15.25%	19.39%	23.45%
67%	0.00%	0.02%	0.47%	1.98%	4.59%	8.02%	11.92%	16.04%	20.18%	24.23%
69%	0.00%	0.03%	0.59%	2.30%	5.11%	8.68%	12.68%	16.83%	20.98%	25.00%
71%	0.00%	0.04%	0.73%	2.64%	5.65%	9.36%	13.44%	17.63%	21.77%	25.77%
73%	0.00%	0.06%	0.88%	3.01%	6.21%	10.05%	14.21%	18.42%	22.56%	26.53%
75%	0.00%	0.08%	1.06%	3.40%	6.79%	10.76%	14.98%	19.22%	23.35%	27.30%
77%	0.00%	0.11%	1.26%	3.82%	7.38%	11.47%	15.76%	20.02%	24.14%	28.06%
79%	0.00%	0.15%	1.48%	4.26%	8.00%	12.20%	16.54%	20.82%	24.93%	28.82%
81%	0.00%	0.19%	1.72%	4.72%	8.63%	12.94%	17.32%	21.61%	25.71%	29.57%
83%	0.00%	0.24%	1.98%	5.21%	9.28%	13.68%	18.11%	22.41%	26.49%	30.33%
85%	0.00%	0.30%	2.26%	5.72%	9.94%	14.43%	18.90%	23.21%	27.27%	31.07%
87%	0.00%	0.38%	2.57%	6.24%	10.62%	15.19%	19.70%	24.00%	28.05%	31.82%
89%	0.00%	0.46%	2.90%	6.79%	11.30%	15.95%	20.49%	24.80%	28.83%	32.56%
91%	0.00%	0.56%	3.25%	7.35%	12.00%	16.72%	21.29%	25.59%	29.60%	33.30%
93%	0.00%	0.67%	3.62%	7.93%	12.71%	17.50%	22.09%	26.38%	30.37%	34.04%
95%	0.00%	0.79%	4.01%	8.53%	13.43%	18.28%	22.88%	27.17%	31.13%	34.77%
97%	0.00%	0.93%	4.43%	9.15%	14.16%	19.06%	23.68%	27.96%	31.90%	35.50%
99%	0.00%	1.09%	4.86%	9.78%	14.90%	19.85%	24.48%	28.75%	32.66%	36.23%
101%	0.00%	1.26%	5.32%	10.42%	15.65%	20.64%	25.28%	29.53%	33.41%	36.95%

Example: Suppose a real option exists that has a \$110 million present value of free cash flows (S), \$100 million in implementation costs (X), 33% volatility, 5% risk-free rate and a 1-year maturity, estimate the real options value of this simple option. The calculated profitability ratio is \$110/\$100 or 10% in-the-money. Using the 1-year table, the option value as a percent of asset is 20.13%, for a 10% profitability ratio and 33% volatility. This means that for the \$110 asset value, the option value is 20.13% of \$110 or \$22.15 million. In addition, if the asset value

	Profitability Ratio (% in-the-money)										
Volatility	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
1%	4.88%	13.52%	20.73%	26.83%	32.06%	36.58%	40.55%	44.05%	47.15%	49.94%	52.44%
3%	4.94%	13.52%	20.73%	26.83%	32.06%	36.58%	40.55%	44.05%	47.15%	49.94%	52.44%
5%	5.28%	13.53%	20.73%	26.83%	32.06%	36.58%	40.55%	44.05%	47.15%	49.94%	52.44%
7%	5.83%	13.57%	20.73%	26.83%	32.06%	36.58%	40.55%	44.05%	47.15%	49.94%	52.44%
9%	6.47%	13.71%	20.74%	26.83%	32.06%	36.58%	40.55%	44.05%	47.15%	49.94%	52.44%
11%	7.15%	13.97%	20.79%	26.83%	32.06%	36.58%	40.55%	44.05%	47.15%	49.94%	52.44%
13%	7.86%	14.32%	20.90%	26.86%	32.06%	36.59%	40.55%	44.05%	47.15%	49.94%	52.44%
15%	8.59%	14.76%	21.08%	26.92%	32.07%	36.59%	40.55%	44.05%	47.15%	49.94%	52.44%
17%	9.33%	15.24%	21.33%	27.02%	32.11%	36.60%	40.55%	44.05%	47.15%	49.94%	52.44%
19%	10.08%	15.78%	21.63%	27.17%	32.18%	36.63%	40.56%	44.05%	47.16%	49.94%	52.44%
21%	10.83%	16.34%	21.99%	27.37%	32.28%	36.67%	40.58%	44.06%	47.16%	49.94%	52.44%
23%	11.58%	16.94%	22.40%	27.62%	32.42%	36.75%	40.62%	44.08%	47.17%	49.94%	52.44%
25%	12.34%	17.55%	22.84%	27.91%	32.60%	36.85%	40.68%	44.11%	47.18%	49.95%	52.45%
27%	13.09%	18.18%	23.31%	28.23%	32.81%	36.99%	40.76%	44.16%	47.21%	49.97%	52.46%
29%	13.85%	18.82%	23.81%	28.60%	33.06%	37.16%	40.87%	44.23%	47.26%	49.99%	52.47%
31%	14.61%	19.47%	24.33%	28.99%	33.35%	37.35%	41.01%	44.32%	47.32%	50.03%	52.50%
33%	15.37%	20.13%	24.87%	29.41%	33.66%	37.58%	41.17%	44.43%	47.39%	50.09%	52.53%
35%	16.13%	20.80%	25.42%	29.85%	34.00%	37.84%	41.36%	44.57%	47.49%	50.16%	52.58%
37%	16.89%	21.47%	25.99%	30.31%	34.36%	38.12%	41.57%	44.73%	47.61%	50.24%	52.65%
39%	17.64%	22.15%	26.57%	30.79%	34.75%	38.42%	41.81%	44.91%	47.75%	50.35%	52.73%
41%	18.40%	22.83%	27.16%	31.28%	35.15%	38.75%	42.07%	45.11%	47.91%	50.47%	52.82%
43%	19.16%	23.52%	27.75%	31.79%	35.57%	39.09%	42.34%	45.34%	48.09%	50.62%	52.94%
45%	19.91%	24.20%	28.36%	32.30%	36.01%	39.46%	42.64%	45.58%	48.29%	50.78%	53.06%
47%	20.67%	24.89%	28.97%	32.83%	36.46%	39.83%	42.96%	45.84%	48.50%	50.95%	53.21%
49%	21.42%	25.58%	29.58%	33.37%	36.92%	40.23%	43.29%	46.12%	48.74%	51.15%	53.37%
51%	22.17%	26.27%	30.20%	33.91%	37.39%	40.63%	43.64%	46.42%	48.99%	51.36%	53.55%
53%	22.92%	26.96%	30.82%	34.46%	37.88%	41.05%	44.00%	46.73%	49.25%	51.58%	53.74%
55%	23.66%	27.65%	31.44%	35.02%	38.37%	41.48%	44.37%	47.05%	49.53%	51.82%	53.95%
57%	24.41%	28.34%	32.07%	35.58%	38.87%	41.92%	44.76%	47.39%	49.82%	52.08%	54.16%
59%	25.15%	29.03%	32.70%	36.15%	39.37%	42.37%	45.15%	47.73%	50.12%	52.34%	54.40%
61%	25.89%	29.72%	33.33%	36.72%	39.88%	42.82%	45.56%	48.09%	50.44%	52.62%	54.64%
63%	26.63%	30.40%	33.96%	37.29%	40.40%	43.29%	45.97%	48.46%	50.76%	52.91%	54.90%
65%	27.37%	31.09%	34.59%	37.87%	40.92%	43.75%	46.39%	48.83%	51.10%	53.21%	55.16%
67%	28.10%	31.78%	35.22%	38.44%	41.44%	44.23%	46.82%	49.22%	51.44%	53.51%	55.44%
69%	28.84%	32.46%	35.86%	39.02%	41.97%	44.71%	47.25%	49.61%	51.80%	53.83%	55.73%
71%	29.57%	33.14%	36.49%	39.61%	42.50%	45.19%	47.69%	50.01%	52.16%	54.16%	56.02%
73%	30.29%	33.83%	37.12%	40.19%	43.04%	45.68%	48.13%	50.41%	52.52%	54.49%	56.32%
75%	31.02%	34.50%	37.75%	40.77%	43.57%	46.17%	48.58%	50.82%	52.90%	54.83%	56.63%
77%	31.74%	35.18%	38.38%	41.35%	44.11%	46.66%	49.03%	51.23%	53.28%	55.18%	56.95%
79%	32.46%	35.86%	39.01%	41.94%	44.65%	47.16%	49.49%	51.65%	53.66%	55.53%	57.27%
81%	33.18%	36.53%	39.64%	42.52%	45.19%	47.66%	49.95%	52.08%	54.05%	55.89%	57.60%
83%	33.89%	37.20%	40.27%	43.11%	45.73%	48.16%	50.41%	52.50%	54.45%	56.25%	57.94%
85%	34.60%	37.87%	40.89%	43.69%	46.27%	48.66%	50.88%	52.93%	54.84%	56.62%	58.28%
87%	35.31%	38.54%	41.52%	44.27%	46.81%	49.17%	51.34%	53.37%	55.25%	56.99%	58.63%
89%	36.02%	39.20%	42.14%	44.85%	47.36%	49.67%	51.81%	53.80%	55.65%	57.37%	58.98%
91%	36.72%	39.87%	42.76%	45.43%	47.90%	50.18%	52.28%	54.24%	56.06%	57.75%	59.33%
93%	37.42%	40.53%	43.38%	46.01%	48.44%	50.68%	52.76%	54.68%	56.47%	58.13%	59.69%
95%	38.11%	41.18%	44.00%	46.59%	48.98%	51.19%	53.23%	55.12%	56.88%	58.52%	60.05%
97%	38.81%	41.84%	44.61%	47.17%	49.52%	51.69%	53.70%	55.57%	57.30%	58.91%	60.41%
99%	39.50%	42.49%	45.23%	47.74%	50.06%	52.20%	54.18%	56.01%	57.71%	59.30%	60.78%
101%	40.18%	43.13%	45.84%	48.32%	50.60%	52.70%	54.65%	56.45%	58.13%	59.69%	61.14%

were \$330 million, then the option value is 20.13% of \$330 million or \$66.44 million as long as the 10% profitability ratio remains the same (implementation cost now becomes \$300 million). The option value as a percentage of asset value does not change as long as the maturity, profitability ratio, and volatility remain constant for these tables.

Real Options Analysis Values (3-year maturity at 5% risk-free rate)

	Profitability Ratio (% in-the-money)									
Volatility	-99%	-90%	-80%	-70%	-60%	-50%	-40%	-30%	-20%	-10%
1%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	4.37%
3%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.19%	4.92%
5%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.03%	1.00%	6.00%
7%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%	0.24%	2.11%	7.23%
9%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.07%	0.74%	3.35%	8.51%
11%	0.00%	0.00%	0.00%	0.00%	0.00%	0.02%	0.26%	1.49%	4.65%	9.81%
13%	0.00%	0.00%	0.00%	0.00%	0.00%	0.08%	0.62%	2.42%	5.99%	11.12%
15%	0.00%	0.00%	0.00%	0.00%	0.02%	0.22%	1.16%	3.47%	7.35%	12.44%
17%	0.00%	0.00%	0.00%	0.00%	0.06%	0.49%	1.86%	4.62%	8.72%	13.76%
19%	0.00%	0.00%	0.00%	0.01%	0.16%	0.88%	2.70%	5.84%	10.10%	15.08%
21%	0.00%	0.00%	0.00%	0.03%	0.33%	1.40%	3.65%	7.10%	11.48%	16.40%
23%	0.00%	0.00%	0.00%	0.08%	0.60%	2.05%	4.68%	8.40%	12.86%	17.72%
25%	0.00%	0.00%	0.01%	0.18%	0.96%	2.80%	5.80%	9.72%	14.25%	19.04%
27%	0.00%	0.00%	0.02%	0.32%	1.43%	3.66%	6.97%	11.07%	15.63%	20.35%
29%	0.00%	0.00%	0.05%	0.54%	1.99%	4.60%	8.18%	12.43%	17.00%	21.65%
31%	0.00%	0.00%	0.11%	0.83%	2.65%	5.61%	9.44%	13.79%	18.37%	22.96%
33%	0.00%	0.00%	0.19%	1.20%	3.40%	6.69%	10.73%	15.17%	19.74%	24.25%
35%	0.00%	0.01%	0.32%	1.65%	4.23%	7.82%	12.04%	16.55%	21.10%	25.54%
37%	0.00%	0.02%	0.50%	2.18%	5.14%	9.01%	13.38%	17.93%	22.46%	26.83%
39%	0.00%	0.04%	0.73%	2.80%	6.11%	10.23%	14.73%	19.31%	23.81%	28.10%
41%	0.00%	0.07%	1.03%	3.49%	7.15%	11.48%	16.09%	20.69%	25.15%	29.37%
43%	0.00%	0.12%	1.38%	4.26%	8.24%	12.77%	17.46%	22.07%	26.49%	30.64%
45%	0.00%	0.19%	1.81%	5.10%	9.38%	14.08%	18.83%	23.44%	27.81%	31.89%
47%	0.00%	0.28%	2.30%	6.00%	10.56%	15.40%	20.21%	24.81%	29.13%	33.14%
49%	0.00%	0.41%	2.87%	6.96%	11.77%	16.75%	21.59%	26.18%	30.44%	34.38%
51%	0.00%	0.58%	3.50%	7.97%	13.02%	18.10%	22.98%	27.53%	31.74%	35.61%
53%	0.00%	0.79%	4.20%	9.04%	14.29%	19.47%	24.36%	28.88%	33.03%	36.82%
55%	0.00%	1.05%	4.96%	10.15%	15.59%	20.84%	25.73%	30.22%	34.31%	38.03%
57%	0.00%	1.36%	5.78%	11.31%	16.91%	22.22%	27.11%	31.55%	35.58%	39.24%
59%	0.00%	1.72%	6.66%	12.49%	18.25%	23.60%	28.47%	32.88%	36.84%	40.43%
61%	0.00%	2.13%	7.59%	13.72%	19.60%	24.98%	29.84%	34.19%	38.09%	41.61%
63%	0.00%	2.60%	8.58%	14.97%	20.96%	26.36%	31.19%	35.49%	39.33%	42.77%
65%	0.01%	3.13%	9.61%	16.25%	22.33%	27.74%	32.54%	36.78%	40.56%	43.93%
67%	0.01%	3.72%	10.69%	17.55%	23.70%	29.12%	33.87%	38.06%	41.77%	45.08%
69%	0.02%	4.36%	11.81%	18.87%	25.08%	30.49%	35.20%	39.33%	42.98%	46.21%
71%	0.03%	5.05%	12.96%	20.20%	26.46%	31.85%	36.52%	40.59%	44.17%	47.34%
73%	0.05%	5.81%	14.15%	21.55%	27.84%	33.21%	37.83%	41.83%	45.35%	48.45%
75%	0.08%	6.61%	15.37%	22.91%	29.23%	34.56%	39.12%	43.06%	46.51%	49.55%
77%	0.11%	7.47%	16.62%	24.28%	30.60%	35.90%	40.40%	44.28%	47.66%	50.64%
79%	0.16%	8.38%	17.90%	25.65%	31.98%	37.23%	41.68%	45.49%	48.80%	51.71%
81%	0.21%	9.33%	19.19%	27.03%	33.35%	38.55%	42.93%	46.68%	49.93%	52.78%
83%	0.29%	10.33%	20.50%	28.41%	34.71%	39.86%	44.18%	47.86%	51.04%	53.83%
85%	0.38%	11.37%	21.83%	29.79%	36.06%	41.16%	45.41%	49.02%	52.14%	54.86%
87%	0.49%	12.45%	23.18%	31.17%	37.41%	42.45%	46.63%	50.17%	53.22%	55.89%
89%	0.62%	13.57%	24.53%	32.55%	38.74%	43.72%	47.83%	51.31%	54.30%	56.90%
91%	0.78%	14.72%	25.90%	33.93%	40.07%	44.98%	49.02%	52.43%	55.35%	57.89%
93%	0.97%	15.91%	27.27%	35.30%	41.38%	46.22%	50.20%	53.54%	56.40%	58.88%
95%	1.19%	17.12%	28.65%	36.66%	42.69%	47.46%	51.36%	54.63%	57.42%	59.85%
97%	1.44%	18.37%	30.03%	38.01%	43.98%	48.67%	52.50%	55.70%	58.44%	60.80%
99%	1.72%	19.63%	31.41%	39.36%	45.25%	49.87%	53.63%	56.77%	59.44%	61.75%
101%	2.04%	20.92%	32.79%	40.70%	46.52%	51.06%	54.74%	57.81%	60.42%	62.67%

	Profitability Ratio (% in-the-money)										
Volatility	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
1%	13.93%	21.75%	28.27%	33.79%	38.52%	42.62%	46.21%	49.37%	52.18%	54.70%	56.96%
3%	13.93%	21.75%	28.27%	33.79%	38.52%	42.62%	46.21%	49.37%	52.18%	54.70%	56.96%
5%	14.06%	21.76%	28.27%	33.79%	38.52%	42.62%	46.21%	49.37%	52.18%	54.70%	56.96%
7%	14.51%	21.84%	28.28%	33.79%	38.52%	42.62%	46.21%	49.37%	52.18%	54.70%	56.96%
9%	15.22%	22.09%	28.35%	33.81%	38.52%	42.62%	46.21%	49.37%	52.18%	54.70%	56.96%
11%	16.10%	22.54%	28.54%	33.88%	38.55%	42.63%	46.21%	49.37%	52.18%	54.70%	56.96%
13%	17.08%	23.14%	28.86%	34.03%	38.62%	42.66%	46.22%	49.38%	52.19%	54.70%	56.96%
15%	18.13%	23.87%	29.32%	34.29%	38.76%	42.73%	46.26%	49.40%	52.19%	54.71%	56.97%
17%	19.22%	24.69%	29.88%	34.66%	38.99%	42.87%	46.34%	49.44%	52.22%	54.72%	56.98%
19%	20.35%	25.57%	30.53%	35.13%	39.31%	43.08%	46.48%	49.53%	52.28%	54.76%	57.00%
21%	21.50%	26.51%	31.26%	35.67%	39.71%	43.37%	46.68%	49.67%	52.38%	54.83%	57.05%
23%	22.67%	27.48%	32.05%	36.28%	40.18%	43.72%	46.95%	49.87%	52.52%	54.93%	57.12%
25%	23.84%	28.49%	32.88%	36.96%	40.71%	44.14%	47.27%	50.12%	52.72%	55.08%	57.24%
27%	25.02%	29.52%	33.75%	37.68%	41.30%	44.62%	47.66%	50.43%	52.96%	55.27%	57.39%
29%	26.21%	30.56%	34.65%	38.44%	41.94%	45.15%	48.09%	50.78%	53.25%	55.51%	57.58%
31%	27.40%	31.62%	35.57%	39.23%	42.61%	45.72%	48.57%	51.19%	53.59%	55.79%	57.82%
33%	28.59%	32.68%	36.51%	40.05%	43.32%	46.33%	49.09%	51.63%	53.97%	56.11%	58.09%
35%	29.78%	33.76%	37.46%	40.89%	44.06%	46.97%	49.65%	52.12%	54.38%	56.48%	58.41%
37%	30.96%	34.84%	38.43%	41.75%	44.82%	47.64%	50.24%	52.63%	54.84%	56.87%	58.75%
39%	32.15%	35.92%	39.41%	42.63%	45.60%	48.34%	50.85%	53.18%	55.32%	57.30%	59.13%
41%	33.33%	37.00%	40.39%	43.52%	46.40%	49.05%	51.49%	53.75%	55.83%	57.75%	59.54%
43%	34.50%	38.08%	41.38%	44.41%	47.21%	49.78%	52.15%	54.34%	56.36%	58.24%	59.97%
45%	35.67%	39.16%	42.37%	45.32%	48.03%	50.53%	52.83%	54.96%	56.92%	58.74%	60.43%
47%	36.83%	40.23%	43.36%	46.22%	48.86%	51.29%	53.52%	55.59%	57.50%	59.26%	60.90%
49%	37.99%	41.31%	44.35%	47.14%	49.70%	52.05%	54.23%	56.23%	58.09%	59.81%	61.40%
51%	39.14%	42.38%	45.34%	48.05%	50.54%	52.83%	54.94%	56.89%	58.69%	60.36%	61.92%
53%	40.28%	43.44%	46.33%	48.97%	51.39%	53.61%	55.66%	57.56%	59.31%	60.93%	62.44%
55%	41.42%	44.50%	47.31%	49.88%	52.24%	54.40%	56.39%	58.23%	59.94%	61.52%	62.99%
57%	42.55%	45.55%	48.29%	50.79%	53.08%	55.19%	57.13%	58.92%	60.57%	62.11%	63.54%
59%	43.66%	46.60%	49.27%	51.70%	53.93%	55.98%	57.87%	59.61%	61.22%	62.71%	64.10%
61%	44.77%	47.64%	50.24%	52.61%	54.78%	56.77%	58.61%	60.30%	61.86%	63.32%	64.67%
63%	45.87%	48.67%	51.21%	53.52%	55.63%	57.57%	59.35%	60.99%	62.52%	63.93%	65.25%
65%	46.96%	49.69%	52.17%	54.42%	56.47%	58.36%	60.09%	61.69%	63.17%	64.55%	65.83%
67%	48.04%	50.71%	53.12%	55.31%	57.31%	59.15%	60.84%	62.39%	63.83%	65.17%	66.41%
69%	49.11%	51.71%	54.06%	56.20%	58.15%	59.94%	61.58%	63.09%	64.49%	65.79%	67.00%
71%	50.17%	52.71%	55.00%	57.08%	58.98%	60.72%	62.32%	63.79%	65.15%	66.42%	67.60%
73%	51.22%	53.70%	55.93%	57.96%	59.81%	61.50%	63.06%	64.49%	65.82%	67.05%	68.19%
75%	52.25%	54.67%	56.85%	58.83%	60.63%	62.28%	63.79%	65.18%	66.47%	67.67%	68.79%
77%	53.28%	55.64%	57.77%	59.69%	61.45%	63.05%	64.52%	65.88%	67.13%	68.30%	69.38%
79%	54.29%	56.60%	58.67%	60.55%	62.25%	63.81%	65.25%	66.57%	67.79%	68.92%	69.97%
81%	55.30%	57.54%	59.57%	61.39%	63.06%	64.57%	65.97%	67.25%	68.44%	69.54%	70.57%
83%	56.29%	58.48%	60.45%	62.23%	63.85%	65.33%	66.68%	67.93%	69.09%	70.16%	71.16%
85%	57.27%	59.41%	61.33%	63.06%	64.64%	66.07%	67.39%	68.61%	69.73%	70.78%	71.75%
87%	58.23%	60.32%	62.19%	63.88%	65.41%	66.81%	68.10%	69.28%	70.37%	71.39%	72.33%
89%	59.19%	61.22%	63.05%	64.69%	66.18%	67.55%	68.80%	69.95%	71.01%	72.00%	72.91%
91%	60.13%	62.11%	63.89%	65.49%	66.95%	68.27%	69.49%	70.61%	71.64%	72.60%	73.49%
93%	61.06%	62.99%	64.72%	66.28%	67.70%	68.99%	70.17%	71.26%	72.27%	73.20%	74.06%
95%	61.97%	63.86%	65.55%	67.07%	68.44%	69.70%	70.85%	71.91%	72.88%	73.79%	74.63%
97%	62.88%	64.72%	66.36%	67.84%	69.18%	70.40%	71.52%	72.55%	73.50%	74.38%	75.20%
99%	63.77%	65.56%	67.16%	68.60%	69.90%	71.09%	72.18%	73.18%	74.10%	74.96%	75.76%
101%	64.65%	66.39%	67.95%	69.35%	70.62%	71.77%	72.83%	73.81%	74.70%	75.54%	76.31%

Real Options Analysis Values (5-year maturity at 5% risk-free rate)

	Profitability Ratio (% in-the-money)									
Volatility	-99%	-90%	-80%	-70%	-60%	-50%	-40%	-30%	-20%	-10%
1%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	2.77%	13.47%
3%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.17%	4.17%	13.50%
5%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.04%	1.07%	5.85%	13.95%
7%	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%	0.35%	2.43%	7.57%	14.86%
9%	0.00%	0.00%	0.00%	0.00%	0.00%	0.12%	1.05%	3.99%	9.30%	16.05%
11%	0.00%	0.00%	0.00%	0.00%	0.04%	0.43%	2.07%	5.66%	11.04%	17.39%
13%	0.00%	0.00%	0.00%	0.01%	0.15%	1.00%	3.32%	7.38%	12.78%	18.81%
15%	0.00%	0.00%	0.00%	0.03%	0.41%	1.80%	4.74%	9.14%	14.51%	20.28%
17%	0.00%	0.00%	0.00%	0.12%	0.84%	2.82%	6.27%	10.91%	16.24%	21.78%
19%	0.00%	0.00%	0.02%	0.29%	1.46%	4.01%	7.88%	12.69%	17.96%	23.31%
21%	0.00%	0.00%	0.05%	0.58%	2.26%	5.34%	9.56%	14.47%	19.67%	24.84%
23%	0.00%	0.00%	0.13%	1.00%	3.22%	6.78%	11.27%	16.26%	21.38%	26.39%
25%	0.00%	0.00%	0.26%	1.57%	4.33%	8.30%	13.01%	18.04%	23.07%	27.93%
27%	0.00%	0.01%	0.48%	2.29%	5.57%	9.90%	14.77%	19.81%	24.76%	29.47%
29%	0.00%	0.04%	0.80%	3.15%	6.92%	11.55%	16.55%	21.58%	26.43%	31.01%
31%	0.00%	0.08%	1.23%	4.14%	8.35%	13.24%	18.33%	23.33%	28.10%	32.54%
33%	0.00%	0.16%	1.77%	5.24%	9.87%	14.96%	20.11%	25.08%	29.75%	34.07%
35%	0.00%	0.28%	2.43%	6.46%	11.44%	16.71%	21.90%	26.81%	31.38%	35.59%
37%	0.00%	0.46%	3.21%	7.77%	13.07%	18.48%	23.67%	28.53%	33.00%	37.09%
39%	0.00%	0.70%	4.11%	9.17%	14.74%	20.25%	25.45%	30.24%	34.61%	38.59%
41%	0.00%	1.03%	5.11%	10.64%	16.45%	22.03%	27.21%	31.93%	36.20%	40.07%
43%	0.00%	1.44%	6.22%	12.18%	18.18%	23.82%	28.96%	33.60%	37.78%	41.54%
45%	0.00%	1.94%	7.42%	13.77%	19.94%	25.60%	30.70%	35.26%	39.34%	42.99%
47%	0.00%	2.53%	8.71%	15.41%	21.71%	27.38%	32.42%	36.90%	40.88%	44.43%
49%	0.01%	3.23%	10.07%	17.09%	23.48%	29.15%	34.13%	38.52%	42.40%	45.86%
51%	0.01%	4.02%	11.51%	18.79%	25.27%	30.91%	35.82%	40.12%	43.91%	47.26%
53%	0.03%	4.90%	13.01%	20.53%	27.05%	32.66%	37.50%	41.71%	45.40%	48.65%
55%	0.05%	5.88%	14.57%	22.28%	28.83%	34.39%	39.15%	43.27%	46.86%	50.03%
57%	0.08%	6.95%	16.17%	24.05%	30.61%	36.11%	40.79%	44.81%	48.31%	51.38%
59%	0.12%	8.10%	17.82%	25.83%	32.38%	37.81%	42.40%	46.33%	49.74%	52.72%
61%	0.19%	9.33%	19.50%	27.61%	34.13%	39.50%	43.99%	47.83%	51.14%	54.04%
63%	0.28%	10.64%	21.21%	29.39%	35.88%	41.16%	45.56%	49.30%	52.52%	55.33%
65%	0.39%	12.02%	22.95%	31.17%	37.61%	42.80%	47.11%	50.76%	53.89%	56.61%
67%	0.55%	13.46%	24.70%	32.95%	39.32%	44.42%	48.64%	52.19%	55.23%	57.87%
69%	0.74%	14.96%	26.47%	34.72%	41.01%	46.02%	50.14%	53.59%	56.55%	59.11%
71%	0.99%	16.51%	28.25%	36.48%	42.69%	47.60%	51.61%	54.97%	57.84%	60.32%
73%	1.28%	18.11%	30.03%	38.22%	44.34%	49.15%	53.06%	56.33%	59.11%	61.52%
75%	1.63%	19.75%	31.81%	39.95%	45.97%	50.67%	54.49%	57.67%	60.36%	62.69%
77%	2.04%	21.42%	33.60%	41.66%	47.58%	52.17%	55.89%	58.97%	61.59%	63.85%
79%	2.52%	23.12%	35.37%	43.35%	49.16%	53.65%	57.26%	60.26%	62.79%	64.98%
81%	3.06%	24.85%	37.14%	45.03%	50.71%	55.10%	58.61%	61.52%	63.97%	66.09%
83%	3.68%	26.60%	38.90%	46.67%	52.25%	56.52%	59.93%	62.75%	65.13%	67.17%
85%	4.37%	28.36%	40.64%	48.30%	53.75%	57.91%	61.23%	63.96%	66.26%	68.24%
87%	5.13%	30.13%	42.37%	49.90%	55.23%	59.27%	62.50%	65.14%	67.37%	69.28%
89%	5.97%	31.92%	44.08%	51.48%	56.68%	60.61%	63.74%	66.30%	68.46%	70.30%
91%	6.88%	33.70%	45.77%	53.02%	58.10%	61.92%	64.95%	67.44%	69.52%	71.30%
93%	7.87%	35.48%	47.43%	54.55%	59.49%	63.20%	66.14%	68.54%	70.55%	72.27%
95%	8.93%	37.26%	49.08%	56.04%	60.85%	64.46%	67.30%	69.62%	71.57%	73.23%
97%	10.06%	39.03%	50.70%	57.50%	62.18%	65.68%	68.44%	70.68%	72.56%	74.16%
99%	11.26%	40.79%	52.29%	58.94%	63.49%	66.88%	69.54%	71.71%	73.52%	75.07%
101%	12.52%	42.54%	53.86%	60.34%	64.76%	68.05%	70.62%	72.72%	74.47%	75.95%

Profitability Ratio (% in-the-money)											
Volatility	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
1%	22.12%	29.20%	35.10%	40.09%	44.37%	48.08%	51.32%	54.19%	56.73%	59.01%	61.06%
3%	22.12%	29.20%	35.10%	40.09%	44.37%	48.08%	51.32%	54.19%	56.73%	59.01%	61.06%
5%	22.16%	29.20%	35.10%	40.09%	44.37%	48.08%	51.32%	54.19%	56.73%	59.01%	61.06%
7%	22.44%	29.26%	35.11%	40.09%	44.37%	48.08%	51.32%	54.19%	56.73%	59.01%	61.06%
9%	23.03%	29.50%	35.19%	40.12%	44.38%	48.08%	51.33%	54.19%	56.73%	59.01%	61.06%
11%	23.86%	29.95%	35.41%	40.22%	44.42%	48.10%	51.33%	54.19%	56.73%	59.01%	61.06%
13%	24.87%	30.60%	35.80%	40.44%	44.55%	48.17%	51.37%	54.21%	56.74%	59.02%	61.06%
15%	26.00%	31.40%	36.35%	40.80%	44.77%	48.31%	51.46%	54.26%	56.78%	59.04%	61.08%
17%	27.21%	32.33%	37.03%	41.29%	45.11%	48.55%	51.62%	54.37%	56.85%	59.09%	61.11%
19%	28.49%	33.35%	37.82%	41.89%	45.57%	48.88%	51.86%	54.56%	56.98%	59.18%	61.18%
21%	29.80%	34.43%	38.70%	42.59%	46.11%	49.31%	52.20%	54.81%	57.18%	59.34%	61.30%
23%	31.14%	35.57%	39.64%	43.36%	46.75%	49.82%	52.61%	55.14%	57.45%	59.55%	61.47%
25%	32.50%	36.75%	40.64%	44.21%	47.45%	50.41%	53.10%	55.55%	57.78%	59.83%	61.70%
27%	33.88%	37.95%	41.69%	45.10%	48.22%	51.06%	53.65%	56.02%	58.18%	60.16%	61.99%
29%	35.26%	39.18%	42.77%	46.05%	49.04%	51.77%	54.26%	56.55%	58.64%	60.56%	62.33%
31%	36.65%	40.42%	43.87%	47.02%	49.89%	52.52%	54.92%	57.13%	59.15%	61.01%	62.72%
33%	38.04%	41.68%	44.99%	48.02%	50.79%	53.31%	55.63%	57.75%	59.70%	61.50%	63.16%
35%	39.43%	42.94%	46.13%	49.05%	51.71%	54.14%	56.37%	58.41%	60.30%	62.03%	63.64%
37%	40.81%	44.20%	47.28%	50.09%	52.65%	54.99%	57.14%	59.11%	60.93%	62.60%	64.15%
39%	42.19%	45.46%	48.43%	51.14%	53.60%	55.86%	57.93%	59.83%	61.58%	63.20%	64.70%
41%	43.56%	46.72%	49.59%	52.20%	54.58%	56.75%	58.75%	60.58%	62.27%	63.83%	65.28%
43%	44.92%	47.98%	50.75%	53.26%	55.56%	57.65%	59.58%	61.34%	62.98%	64.48%	65.88%
45%	46.27%	49.23%	51.90%	54.33%	56.54%	58.57%	60.42%	62.13%	63.70%	65.15%	66.50%
47%	47.61%	50.47%	53.06%	55.40%	57.53%	59.48%	61.27%	62.92%	64.44%	65.84%	67.14%
49%	48.94%	51.71%	54.20%	56.47%	58.53%	60.41%	62.13%	63.72%	65.19%	66.54%	67.80%
51%	50.25%	52.93%	55.34%	57.53%	59.52%	61.33%	63.00%	64.53%	65.94%	67.25%	68.47%
53%	51.55%	54.14%	56.48%	58.59%	60.51%	62.26%	63.87%	65.34%	66.71%	67.97%	69.14%
55%	52.84%	55.34%	57.60%	59.64%	61.49%	63.18%	64.73%	66.16%	67.48%	68.70%	69.83%
57%	54.10%	56.53%	58.71%	60.68%	62.47%	64.11%	65.60%	66.98%	68.25%	69.42%	70.52%
59%	55.36%	57.71%	59.82%	61.72%	63.45%	65.02%	66.47%	67.79%	69.02%	70.15%	71.21%
61%	56.59%	58.87%	60.90%	62.74%	64.41%	65.93%	67.33%	68.61%	69.79%	70.89%	71.90%
63%	57.81%	60.01%	61.98%	63.76%	65.37%	66.84%	68.18%	69.42%	70.56%	71.62%	72.60%
65%	59.01%	61.14%	63.04%	64.76%	66.32%	67.73%	69.03%	70.22%	71.32%	72.34%	73.29%
67%	60.19%	62.25%	64.09%	65.75%	67.25%	68.62%	69.87%	71.02%	72.09%	73.07%	73.98%
69%	61.36%	63.35%	65.13%	66.73%	68.18%	69.50%	70.71%	71.82%	72.84%	73.79%	74.67%
71%	62.50%	64.43%	66.15%	67.69%	69.09%	70.37%	71.53%	72.60%	73.59%	74.50%	75.35%
73%	63.63%	65.49%	67.15%	68.64%	69.99%	71.22%	72.35%	73.38%	74.33%	75.21%	76.03%
75%	64.73%	66.53%	68.13%	69.58%	70.88%	72.07%	73.15%	74.15%	75.06%	75.91%	76.70%
77%	65.82%	67.56%	69.10%	70.50%	71.75%	72.90%	73.94%	74.90%	75.79%	76.61%	77.37%
79%	66.88%	68.56%	70.06%	71.40%	72.61%	73.72%	74.72%	75.65%	76.50%	77.29%	78.03%
81%	67.93%	69.55%	70.99%	72.29%	73.46%	74.52%	75.49%	76.39%	77.21%	77.97%	78.68%
83%	68.95%	70.52%	71.91%	73.16%	74.29%	75.31%	76.25%	77.11%	77.90%	78.64%	79.32%
85%	69.96%	71.47%	72.81%	74.02%	75.10%	76.09%	76.99%	77.82%	78.59%	79.29%	79.95%
87%	70.94%	72.40%	73.69%	74.86%	75.90%	76.86%	77.73%	78.52%	79.26%	79.94%	80.57%
89%	71.90%	73.31%	74.56%	75.68%	76.69%	77.61%	78.44%	79.21%	79.92%	80.57%	81.18%
91%	72.84%	74.20%	75.41%	76.48%	77.46%	78.34%	79.15%	79.89%	80.57%	81.20%	81.78%
93%	73.76%	75.07%	76.23%	77.27%	78.21%	79.06%	79.84%	80.55%	81.20%	81.81%	82.37%
95%	74.66%	75.93%	77.04%	78.05%	78.95%	79.77%	80.51%	81.20%	81.83%	82.41%	82.95%
97%	75.54%	76.76%	77.84%	78.80%	79.67%	80.46%	81.17%	81.83%	82.44%	83.00%	83.52%
99%	76.40%	77.57%	78.61%	79.54%	80.37%	81.13%	81.82%	82.45%	83.04%	83.58%	84.08%
101%	77.24%	78.37%	79.36%	80.26%	81.06%	81.79%	82.45%	83.06%	83.62%	84.14%	84.62%

Real Options Analysis Values (7-year maturity at 5% risk-free rate)

	Profitability Ratio (% in-the-money)									
Volatility	-99%	-90%	-80%	-70%	-60%	-50%	-40%	-30%	-20%	-10%
1%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.76%	11.91%	21.70%
3%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.07%	2.85%	12.09%	21.70%
5%	0.00%	0.00%	0.00%	0.00%	0.00%	0.02%	0.78%	4.96%	13.03%	21.85%
7%	0.00%	0.00%	0.00%	0.00%	0.01%	0.27%	2.13%	7.07%	14.46%	22.41%
9%	0.00%	0.00%	0.00%	0.00%	0.09%	0.94%	3.83%	9.18%	16.10%	23.36%
11%	0.00%	0.00%	0.00%	0.02%	0.38%	2.01%	5.71%	11.28%	17.85%	24.56%
13%	0.00%	0.00%	0.00%	0.11%	0.94%	3.39%	7.70%	13.37%	19.65%	25.93%
15%	0.00%	0.00%	0.01%	0.34%	1.79%	4.99%	9.74%	15.45%	21.49%	27.41%
17%	0.00%	0.00%	0.06%	0.75%	2.91%	6.75%	11.82%	17.52%	23.34%	28.95%
19%	0.00%	0.00%	0.18%	1.38%	4.25%	8.63%	13.92%	19.58%	25.20%	30.54%
21%	0.00%	0.01%	0.41%	2.23%	5.78%	10.59%	16.03%	21.62%	27.06%	32.16%
23%	0.00%	0.03%	0.77%	3.29%	7.45%	12.61%	18.14%	23.65%	28.91%	33.80%
25%	0.00%	0.08%	1.30%	4.53%	9.24%	14.66%	20.25%	25.67%	30.76%	35.45%
27%	0.00%	0.17%	2.00%	5.94%	11.12%	16.75%	22.35%	27.66%	32.59%	37.10%
29%	0.00%	0.33%	2.87%	7.49%	13.07%	18.85%	24.43%	29.64%	34.42%	38.75%
31%	0.00%	0.59%	3.91%	9.17%	15.08%	20.96%	26.50%	31.60%	36.22%	40.40%
33%	0.00%	0.94%	5.11%	10.95%	17.13%	23.07%	28.56%	33.54%	38.02%	42.04%
35%	0.00%	1.43%	6.45%	12.81%	19.20%	25.18%	30.59%	35.45%	39.79%	43.67%
37%	0.00%	2.04%	7.93%	14.74%	21.30%	27.27%	32.61%	37.34%	41.54%	45.28%
39%	0.00%	2.79%	9.52%	16.73%	23.41%	29.36%	34.60%	39.21%	43.27%	46.88%
41%	0.01%	3.68%	11.22%	18.76%	25.52%	31.42%	36.56%	41.05%	44.98%	48.46%
43%	0.02%	4.70%	13.00%	20.82%	27.63%	33.47%	38.50%	42.86%	46.67%	50.02%
45%	0.04%	5.86%	14.87%	22.91%	29.73%	35.50%	40.42%	44.65%	48.33%	51.56%
47%	0.07%	7.14%	16.79%	25.01%	31.82%	37.50%	42.30%	46.41%	49.97%	53.08%
49%	0.12%	8.55%	18.77%	27.12%	33.89%	39.47%	44.16%	48.14%	51.58%	54.58%
51%	0.20%	10.06%	20.80%	29.23%	35.95%	41.42%	45.98%	49.84%	53.17%	56.06%
53%	0.32%	11.68%	22.85%	31.34%	37.98%	43.34%	47.77%	51.52%	54.72%	57.51%
55%	0.48%	13.38%	24.94%	33.43%	39.98%	45.22%	49.53%	53.16%	56.25%	58.94%
57%	0.70%	15.17%	27.03%	35.52%	41.96%	47.08%	51.26%	54.77%	57.75%	60.34%
59%	0.99%	17.03%	29.14%	37.58%	43.91%	48.90%	52.96%	56.34%	59.22%	61.71%
61%	1.34%	18.95%	31.25%	39.62%	45.83%	50.69%	54.62%	57.89%	60.67%	63.06%
63%	1.78%	20.92%	33.36%	41.64%	47.72%	52.44%	56.25%	59.40%	62.08%	64.38%
65%	2.31%	22.93%	35.46%	43.63%	49.57%	54.15%	57.84%	60.88%	63.46%	65.67%
67%	2.93%	24.98%	37.55%	45.59%	51.38%	55.83%	59.39%	62.33%	64.81%	66.93%
69%	3.65%	27.06%	39.62%	47.52%	53.16%	57.47%	60.91%	63.74%	66.13%	68.17%
71%	4.47%	29.15%	41.66%	49.42%	54.91%	59.08%	62.40%	65.12%	67.41%	69.37%
73%	5.40%	31.25%	43.69%	51.28%	56.61%	60.65%	63.85%	66.47%	68.67%	70.55%
75%	6.44%	33.36%	45.68%	53.10%	58.27%	62.17%	65.26%	67.78%	69.89%	71.70%
77%	7.58%	35.47%	47.65%	54.88%	59.90%	63.66%	66.63%	69.06%	71.09%	72.82%
79%	8.82%	37.58%	49.58%	56.62%	61.48%	65.12%	67.97%	70.30%	72.25%	73.91%
81%	10.16%	39.67%	51.47%	58.33%	63.03%	66.53%	69.28%	71.51%	73.38%	74.97%
83%	11.60%	41.74%	53.33%	59.99%	64.53%	67.90%	70.54%	72.69%	74.48%	76.00%
85%	13.13%	43.80%	55.15%	61.61%	65.99%	69.24%	71.77%	73.83%	75.54%	77.00%
87%	14.74%	45.83%	56.93%	63.19%	67.42%	70.54%	72.97%	74.94%	76.58%	77.97%
89%	16.43%	47.83%	58.67%	64.73%	68.80%	71.79%	74.13%	76.01%	77.58%	78.91%
91%	18.20%	49.80%	60.37%	66.22%	70.14%	73.02%	75.25%	77.06%	78.56%	79.83%
93%	20.03%	51.74%	62.02%	67.67%	71.44%	74.20%	76.34%	78.07%	79.50%	80.71%
95%	21.93%	53.64%	63.63%	69.08%	72.70%	75.34%	77.39%	79.05%	80.41%	81.57%
97%	23.87%	55.50%	65.19%	70.44%	73.92%	76.45%	78.41%	79.99%	81.30%	82.40%
99%	25.86%	57.32%	66.71%	71.77%	75.10%	77.52%	79.40%	80.90%	82.15%	83.20%
101%	27.88%	59.10%	68.19%	73.05%	76.24%	78.56%	80.35%	81.79%	82.98%	83.98%

	Profitability Ratio (% in-the-money)										
Volatility	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
1%	29.53%	35.94%	41.28%	45.79%	49.67%	53.02%	55.96%	58.55%	60.85%	62.91%	64.77%
3%	29.53%	35.94%	41.28%	45.79%	49.67%	53.02%	55.96%	58.55%	60.85%	62.91%	64.77%
5%	29.55%	35.94%	41.28%	45.79%	49.67%	53.02%	55.96%	58.55%	60.85%	62.91%	64.77%
7%	29.71%	35.98%	41.28%	45.79%	49.67%	53.02%	55.96%	58.55%	60.85%	62.91%	64.77%
9%	30.16%	36.16%	41.36%	45.82%	49.67%	53.02%	55.96%	58.55%	60.85%	62.91%	64.77%
11%	30.88%	36.57%	41.57%	45.93%	49.73%	53.05%	55.97%	58.55%	60.85%	62.91%	64.77%
13%	31.84%	37.19%	41.96%	46.17%	49.87%	53.14%	56.02%	58.58%	60.87%	62.92%	64.77%
15%	32.95%	38.00%	42.53%	46.56%	50.14%	53.31%	56.14%	58.66%	60.92%	62.96%	64.80%
17%	34.18%	38.95%	43.24%	47.09%	50.53%	53.60%	56.35%	58.82%	61.04%	63.04%	64.86%
19%	35.49%	40.01%	44.08%	47.75%	51.04%	54.00%	56.66%	59.06%	61.22%	63.18%	64.97%
21%	36.87%	41.15%	45.02%	48.52%	51.67%	54.51%	57.07%	59.38%	61.49%	63.40%	65.14%
23%	38.28%	42.36%	46.04%	49.37%	52.38%	55.10%	57.56%	59.80%	61.83%	63.69%	65.38%
25%	39.73%	43.61%	47.13%	50.30%	53.18%	55.78%	58.14%	60.29%	62.25%	64.05%	65.69%
27%	41.20%	44.90%	48.26%	51.29%	54.03%	56.53%	58.79%	60.86%	62.75%	64.48%	66.06%
29%	42.68%	46.22%	49.42%	52.32%	54.94%	57.33%	59.50%	61.48%	63.30%	64.96%	66.50%
31%	44.16%	47.55%	50.62%	53.38%	55.90%	58.18%	60.26%	62.16%	63.91%	65.51%	66.99%
33%	45.65%	48.90%	51.83%	54.48%	56.88%	59.07%	61.06%	62.88%	64.56%	66.10%	67.52%
35%	47.14%	50.25%	53.06%	55.59%	57.89%	59.98%	61.90%	63.65%	65.25%	66.73%	68.10%
37%	48.61%	51.60%	54.29%	56.72%	58.92%	60.93%	62.76%	64.44%	65.98%	67.40%	68.71%
39%	50.08%	52.95%	55.53%	57.86%	59.97%	61.89%	63.64%	65.25%	66.73%	68.10%	69.36%
41%	51.54%	54.30%	56.77%	59.00%	61.02%	62.86%	64.55%	66.09%	67.51%	68.82%	70.03%
43%	52.99%	55.63%	58.01%	60.15%	62.08%	63.85%	65.46%	66.94%	68.30%	69.56%	70.72%
45%	54.42%	56.96%	59.24%	61.29%	63.15%	64.84%	66.38%	67.80%	69.11%	70.31%	71.43%
47%	55.83%	58.27%	60.46%	62.43%	64.21%	65.83%	67.31%	68.67%	69.92%	71.08%	72.15%
49%	57.22%	59.57%	61.67%	63.56%	65.27%	66.82%	68.24%	69.54%	70.74%	71.85%	72.88%
51%	58.60%	60.85%	62.87%	64.68%	66.32%	67.81%	69.17%	70.42%	71.57%	72.63%	73.62%
53%	59.95%	62.12%	64.05%	65.79%	67.36%	68.79%	70.09%	71.29%	72.39%	73.41%	74.36%
55%	61.29%	63.37%	65.22%	66.89%	68.39%	69.76%	71.01%	72.16%	73.22%	74.20%	75.10%
57%	62.60%	64.59%	66.37%	67.97%	69.42%	70.73%	71.93%	73.03%	74.04%	74.98%	75.85%
59%	63.88%	65.80%	67.51%	69.04%	70.43%	71.68%	72.83%	73.89%	74.86%	75.75%	76.59%
61%	65.14%	66.99%	68.62%	70.09%	71.42%	72.63%	73.73%	74.74%	75.67%	76.53%	77.32%
63%	66.38%	68.15%	69.72%	71.13%	72.40%	73.56%	74.61%	75.58%	76.47%	77.29%	78.05%
65%	67.59%	69.29%	70.80%	72.15%	73.37%	74.47%	75.48%	76.41%	77.26%	78.05%	78.78%
67%	68.78%	70.41%	71.85%	73.15%	74.31%	75.37%	76.34%	77.22%	78.04%	78.80%	79.50%
69%	69.94%	71.50%	72.88%	74.12%	75.24%	76.26%	77.18%	78.03%	78.81%	79.53%	80.20%
71%	71.07%	72.57%	73.90%	75.08%	76.15%	77.12%	78.01%	78.82%	79.57%	80.26%	80.90%
73%	72.18%	73.61%	74.88%	76.02%	77.04%	77.97%	78.82%	79.60%	80.31%	80.97%	81.58%
75%	73.26%	74.63%	75.85%	76.94%	77.92%	78.80%	79.61%	80.36%	81.04%	81.67%	82.26%
77%	74.31%	75.63%	76.79%	77.83%	78.77%	79.62%	80.39%	81.10%	81.75%	82.36%	82.92%
79%	75.34%	76.60%	77.71%	78.71%	79.60%	80.41%	81.15%	81.83%	82.45%	83.03%	83.57%
81%	76.34%	77.54%	78.61%	79.56%	80.41%	81.19%	81.89%	82.54%	83.14%	83.69%	84.20%
83%	77.31%	78.46%	79.48%	80.39%	81.20%	81.94%	82.62%	83.24%	83.81%	84.33%	84.82%
85%	78.25%	79.35%	80.33%	81.19%	81.97%	82.68%	83.32%	83.91%	84.46%	84.96%	85.42%
87%	79.17%	80.22%	81.15%	81.98%	82.72%	83.40%	84.01%	84.58%	85.09%	85.57%	86.01%
89%	80.06%	81.06%	81.95%	82.74%	83.45%	84.09%	84.68%	85.22%	85.71%	86.17%	86.59%
91%	80.92%	81.88%	82.73%	83.48%	84.16%	84.77%	85.33%	85.84%	86.31%	86.75%	87.15%
93%	81.76%	82.67%	83.48%	84.20%	84.84%	85.43%	85.96%	86.45%	86.90%	87.31%	87.70%
95%	82.57%	83.44%	84.21%	84.89%	85.51%	86.07%	86.57%	87.04%	87.47%	87.86%	88.23%
97%	83.35%	84.18%	84.92%	85.57%	86.15%	86.69%	87.17%	87.61%	88.02%	88.39%	88.74%
99%	84.11%	84.90%	85.60%	86.22%	86.78%	87.28%	87.74%	88.16%	88.55%	88.91%	89.24%
101%	84.84%	85.60%	86.26%	86.85%	87.38%	87.86%	88.30%	88.70%	89.07%	89.41%	89.72%

Real Options Analysis Values (10-year maturity at 5% risk-free rate)

	Profitability Ratio (% in-the-money)									
Volatility	-99%	-90%	-80%	-70%	-60%	-50%	-40%	-30%	-20%	-10%
1%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.80%	13.35%	24.18%	32.61%
3%	0.00%	0.00%	0.00%	0.00%	0.00%	0.08%	3.28%	13.61%	24.19%	32.61%
5%	0.00%	0.00%	0.00%	0.00%	0.03%	0.93%	5.81%	14.81%	24.41%	32.63%
7%	0.00%	0.00%	0.00%	0.01%	0.32%	2.56%	8.33%	16.56%	25.15%	32.88%
9%	0.00%	0.00%	0.00%	0.09%	1.11%	4.62%	10.84%	18.54%	26.34%	33.48%
11%	0.00%	0.00%	0.01%	0.39%	2.39%	6.89%	13.35%	20.63%	27.82%	34.42%
13%	0.00%	0.00%	0.07%	1.02%	4.06%	9.29%	15.84%	22.77%	29.47%	35.60%
15%	0.00%	0.00%	0.26%	2.01%	6.01%	11.75%	18.31%	24.95%	31.23%	36.96%
17%	0.00%	0.01%	0.65%	3.33%	8.15%	14.25%	20.77%	27.13%	33.06%	38.44%
19%	0.00%	0.05%	1.28%	4.93%	10.43%	16.77%	23.20%	29.32%	34.94%	40.00%
21%	0.00%	0.16%	2.18%	6.76%	12.81%	19.29%	25.61%	31.50%	36.83%	41.62%
23%	0.00%	0.36%	3.34%	8.78%	15.25%	21.81%	28.00%	33.66%	38.74%	43.28%
25%	0.00%	0.71%	4.74%	10.95%	17.74%	24.31%	30.36%	35.81%	40.66%	44.96%
27%	0.00%	1.23%	6.37%	13.24%	20.25%	26.80%	32.69%	37.93%	42.56%	46.66%
29%	0.00%	1.94%	8.18%	15.61%	22.77%	29.26%	34.99%	40.03%	44.46%	48.36%
31%	0.00%	2.85%	10.16%	18.04%	25.30%	31.69%	37.26%	42.11%	46.34%	50.06%
33%	0.01%	3.97%	12.28%	20.52%	27.81%	34.09%	39.49%	44.15%	48.20%	51.74%
35%	0.02%	5.29%	14.51%	23.03%	30.30%	36.46%	41.69%	46.17%	50.04%	53.42%
37%	0.04%	6.79%	16.84%	25.54%	32.77%	38.79%	43.84%	48.15%	51.86%	55.08%
39%	0.09%	8.48%	19.23%	28.07%	35.21%	41.08%	45.96%	50.10%	53.64%	56.72%
41%	0.18%	10.32%	21.68%	30.58%	37.62%	43.32%	48.04%	52.01%	55.40%	58.33%
43%	0.31%	12.30%	24.16%	33.08%	40.00%	45.53%	50.07%	53.88%	57.12%	59.92%
45%	0.51%	14.41%	26.67%	35.56%	42.33%	47.69%	52.07%	55.71%	58.81%	61.48%
47%	0.79%	16.62%	29.20%	38.01%	44.62%	49.81%	54.01%	57.51%	60.47%	63.02%
49%	1.17%	18.92%	31.72%	40.43%	46.86%	51.87%	55.91%	59.26%	62.09%	64.52%
51%	1.67%	21.30%	34.23%	42.80%	49.06%	53.89%	57.77%	60.97%	63.67%	65.99%
53%	2.29%	23.73%	36.73%	45.14%	51.20%	55.86%	59.58%	62.64%	65.22%	67.42%
55%	3.05%	26.20%	39.20%	47.43%	53.30%	57.78%	61.34%	64.27%	66.72%	68.82%
57%	3.95%	28.70%	41.64%	49.67%	55.34%	59.64%	63.06%	65.85%	68.19%	70.19%
59%	5.01%	31.22%	44.04%	51.85%	57.33%	61.46%	64.72%	67.39%	69.62%	71.52%
61%	6.21%	33.74%	46.40%	53.99%	59.26%	63.22%	66.34%	68.88%	71.00%	72.81%
63%	7.57%	36.26%	48.72%	56.07%	61.14%	64.93%	67.91%	70.33%	72.35%	74.06%
65%	9.07%	38.76%	50.98%	58.09%	62.96%	66.59%	69.43%	71.73%	73.65%	75.28%
67%	10.72%	41.24%	53.19%	60.06%	64.73%	68.19%	70.90%	73.09%	74.91%	76.46%
69%	12.50%	43.70%	55.35%	61.97%	66.44%	69.75%	72.32%	74.41%	76.14%	77.61%
71%	14.41%	46.11%	57.45%	63.82%	68.10%	71.25%	73.70%	75.68%	77.32%	78.71%
73%	16.43%	48.49%	59.49%	65.60%	69.69%	72.70%	75.02%	76.90%	78.46%	79.78%
75%	18.56%	50.82%	61.47%	67.33%	71.24%	74.09%	76.30%	78.09%	79.56%	80.81%
77%	20.78%	53.10%	63.39%	69.00%	72.72%	75.44%	77.54%	79.22%	80.62%	81.80%
79%	23.08%	55.32%	65.25%	70.61%	74.16%	76.73%	78.72%	80.32%	81.64%	82.76%
81%	25.44%	57.49%	67.04%	72.16%	75.53%	77.98%	79.86%	81.37%	82.62%	83.68%
83%	27.86%	59.59%	68.77%	73.66%	76.86%	79.17%	80.96%	82.39%	83.56%	84.56%
85%	30.33%	61.64%	70.44%	75.09%	78.13%	80.32%	82.01%	83.36%	84.47%	85.41%
87%	32.82%	63.61%	72.04%	76.47%	79.35%	81.42%	83.01%	84.29%	85.34%	86.22%
89%	35.34%	65.53%	73.58%	77.79%	80.51%	82.47%	83.98%	85.18%	86.17%	87.00%
91%	37.86%	67.38%	75.06%	79.05%	81.63%	83.48%	84.90%	86.03%	86.96%	87.74%
93%	40.38%	69.16%	76.48%	80.26%	82.70%	84.44%	85.78%	86.84%	87.72%	88.46%
95%	42.89%	70.87%	77.84%	81.42%	83.72%	85.36%	86.62%	87.62%	88.45%	89.14%
97%	45.38%	72.52%	79.14%	82.52%	84.69%	86.24%	87.42%	88.36%	89.14%	89.79%
99%	47.84%	74.10%	80.38%	83.57%	85.62%	87.07%	88.19%	89.07%	89.80%	90.40%
101%	50.26%	75.61%	81.56%	84.58%	86.50%	87.87%	88.91%	89.74%	90.42%	90.99%

Volatility	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
	Profitability Ratio (% in-the-money)										
1%	39.35%	44.86%	49.46%	53.34%	56.68%	59.56%	62.09%	64.32%	66.30%	68.08%	69.67%
3%	39.35%	44.86%	49.46%	53.34%	56.68%	59.56%	62.09%	64.32%	66.30%	68.08%	69.67%
5%	39.35%	44.86%	49.46%	53.34%	56.68%	59.56%	62.09%	64.32%	66.30%	68.08%	69.67%
7%	39.42%	44.88%	49.46%	53.34%	56.68%	59.56%	62.09%	64.32%	66.30%	68.08%	69.67%
9%	39.70%	45.00%	49.51%	53.37%	56.69%	59.57%	62.09%	64.32%	66.30%	68.08%	69.67%
11%	40.25%	45.32%	49.69%	53.46%	56.74%	59.60%	62.11%	64.33%	66.31%	68.08%	69.67%
13%	41.06%	45.85%	50.04%	53.69%	56.88%	59.69%	62.17%	64.37%	66.33%	68.10%	69.69%
15%	42.07%	46.59%	50.57%	54.07%	57.15%	59.88%	62.31%	64.47%	66.40%	68.15%	69.72%
17%	43.24%	47.50%	51.27%	54.60%	57.56%	60.20%	62.55%	64.65%	66.55%	68.26%	69.81%
19%	44.52%	48.54%	52.11%	55.28%	58.11%	60.63%	62.90%	64.93%	66.77%	68.44%	69.96%
21%	45.89%	49.68%	53.06%	56.07%	58.76%	61.18%	63.35%	65.31%	67.09%	68.71%	70.18%
23%	47.31%	50.90%	54.10%	56.96%	59.52%	61.83%	63.90%	65.79%	67.50%	69.05%	70.48%
25%	48.78%	52.18%	55.22%	57.93%	60.36%	62.56%	64.54%	66.34%	67.98%	69.48%	70.85%
27%	50.28%	53.51%	56.38%	58.96%	61.27%	63.36%	65.25%	66.97%	68.54%	69.98%	71.30%
29%	51.80%	54.86%	57.59%	60.04%	62.24%	64.22%	66.03%	67.67%	69.17%	70.54%	71.81%
31%	53.33%	56.24%	58.83%	61.15%	63.24%	65.13%	66.85%	68.41%	69.85%	71.16%	72.37%
33%	54.86%	57.62%	60.09%	62.29%	64.28%	66.08%	67.71%	69.20%	70.57%	71.82%	72.98%
35%	56.39%	59.01%	61.35%	63.45%	65.34%	67.05%	68.61%	70.03%	71.33%	72.52%	73.63%
37%	57.90%	60.40%	62.63%	64.62%	66.42%	68.04%	69.52%	70.88%	72.12%	73.26%	74.31%
39%	59.41%	61.79%	63.90%	65.80%	67.51%	69.05%	70.46%	71.75%	72.93%	74.02%	75.02%
41%	60.90%	63.16%	65.17%	66.98%	68.60%	70.07%	71.41%	72.63%	73.76%	74.79%	75.75%
43%	62.37%	64.52%	66.43%	68.15%	69.69%	71.09%	72.36%	73.53%	74.60%	75.58%	76.49%
45%	63.81%	65.86%	67.68%	69.31%	70.78%	72.11%	73.32%	74.43%	75.44%	76.38%	77.25%
47%	65.23%	67.18%	68.92%	70.46%	71.86%	73.12%	74.27%	75.33%	76.29%	77.18%	78.01%
49%	66.63%	68.48%	70.13%	71.60%	72.93%	74.13%	75.22%	76.22%	77.14%	77.99%	78.77%
51%	68.00%	69.76%	71.33%	72.73%	73.99%	75.13%	76.16%	77.11%	77.99%	78.79%	79.54%
53%	69.33%	71.01%	72.50%	73.83%	75.03%	76.11%	77.10%	78.00%	78.83%	79.59%	80.30%
55%	70.64%	72.24%	73.65%	74.91%	76.05%	77.08%	78.01%	78.87%	79.66%	80.38%	81.06%
57%	71.92%	73.43%	74.78%	75.97%	77.05%	78.03%	78.92%	79.73%	80.48%	81.17%	81.80%
59%	73.16%	74.60%	75.88%	77.01%	78.04%	78.96%	79.80%	80.57%	81.28%	81.94%	82.54%
61%	74.37%	75.74%	76.95%	78.03%	79.00%	79.87%	80.67%	81.40%	82.08%	82.70%	83.27%
63%	75.55%	76.84%	77.99%	79.01%	79.93%	80.77%	81.52%	82.22%	82.85%	83.44%	83.98%
65%	76.69%	77.92%	79.01%	79.98%	80.85%	81.64%	82.35%	83.01%	83.61%	84.17%	84.68%
67%	77.80%	78.96%	79.99%	80.91%	81.74%	82.48%	83.16%	83.78%	84.35%	84.88%	85.37%
69%	78.87%	79.98%	80.95%	81.82%	82.60%	83.31%	83.95%	84.54%	85.08%	85.57%	86.04%
71%	79.91%	80.95%	81.88%	82.70%	83.44%	84.11%	84.71%	85.27%	85.78%	86.25%	86.69%
73%	80.91%	81.90%	82.77%	83.55%	84.25%	84.88%	85.46%	85.98%	86.46%	86.91%	87.32%
75%	81.88%	82.82%	83.64%	84.38%	85.04%	85.63%	86.18%	86.67%	87.13%	87.55%	87.94%
77%	82.81%	83.70%	84.48%	85.17%	85.80%	86.36%	86.87%	87.34%	87.77%	88.17%	88.53%
79%	83.71%	84.55%	85.29%	85.94%	86.53%	87.06%	87.54%	87.98%	88.39%	88.76%	89.11%
81%	84.58%	85.37%	86.06%	86.68%	87.24%	87.74%	88.19%	88.61%	88.99%	89.34%	89.67%
83%	85.41%	86.16%	86.81%	87.39%	87.92%	88.39%	88.82%	89.21%	89.57%	89.90%	90.21%
85%	86.21%	86.91%	87.53%	88.08%	88.57%	89.01%	89.42%	89.79%	90.13%	90.44%	90.73%
87%	86.98%	87.64%	88.22%	88.74%	89.20%	89.62%	90.00%	90.34%	90.66%	90.96%	91.23%
89%	87.71%	88.33%	88.88%	89.37%	89.80%	90.19%	90.55%	90.88%	91.18%	91.45%	91.71%
91%	88.42%	89.00%	89.51%	89.97%	90.38%	90.75%	91.08%	91.39%	91.67%	91.93%	92.17%
93%	89.09%	89.64%	90.12%	90.55%	90.93%	91.28%	91.59%	91.88%	92.15%	92.39%	92.62%
95%	89.73%	90.25%	90.70%	91.10%	91.46%	91.79%	92.08%	92.35%	92.60%	92.83%	93.04%
97%	90.34%	90.83%	91.25%	91.63%	91.97%	92.27%	92.55%	92.80%	93.03%	93.25%	93.45%
99%	90.93%	91.38%	91.78%	92.13%	92.45%	92.73%	92.99%	93.23%	93.45%	93.65%	93.84%
101%	91.48%	91.91%	92.28%	92.61%	92.91%	93.18%	93.42%	93.64%	93.85%	94.03%	94.21%

Real Options Analysis Values (15-year maturity at 5% risk-free rate)

	Profitability Ratio (% in-the-money)									
Volatility	-99%	-90%	-80%	-70%	-60%	-50%	-40%	-30%	-20%	-10%
1%	0.00%	0.00%	0.00%	0.00%	0.00%	5.65%	21.27%	32.52%	40.95%	47.51%
3%	0.00%	0.00%	0.00%	0.00%	0.43%	7.80%	21.35%	32.52%	40.95%	47.51%
5%	0.00%	0.00%	0.00%	0.08%	2.27%	10.58%	22.17%	32.64%	40.97%	47.52%
7%	0.00%	0.00%	0.01%	0.66%	4.84%	13.48%	23.75%	33.24%	41.16%	47.57%
9%	0.00%	0.00%	0.12%	1.96%	7.70%	16.39%	25.76%	34.36%	41.71%	47.83%
11%	0.00%	0.00%	0.52%	3.87%	10.71%	19.31%	27.98%	35.84%	42.63%	48.38%
13%	0.00%	0.03%	1.36%	6.20%	13.77%	22.21%	30.31%	37.56%	43.84%	49.21%
15%	0.00%	0.14%	2.66%	8.83%	16.86%	25.09%	32.71%	39.44%	45.27%	50.27%
17%	0.00%	0.42%	4.39%	11.66%	19.95%	27.94%	35.13%	41.42%	46.84%	51.51%
19%	0.00%	0.95%	6.47%	14.61%	23.02%	30.76%	37.57%	43.45%	48.52%	52.88%
21%	0.00%	1.79%	8.85%	17.63%	26.06%	33.54%	39.99%	45.52%	50.27%	54.35%
23%	0.00%	2.96%	11.46%	20.71%	29.07%	36.27%	42.40%	47.61%	52.06%	55.89%
25%	0.00%	4.45%	14.24%	23.79%	32.04%	38.96%	44.78%	49.69%	53.87%	57.47%
27%	0.02%	6.25%	17.14%	26.88%	34.95%	41.61%	47.13%	51.76%	55.70%	59.07%
29%	0.05%	8.33%	20.14%	29.95%	37.82%	44.19%	49.44%	53.81%	57.52%	60.69%
31%	0.13%	10.65%	23.19%	32.98%	40.63%	46.73%	51.70%	55.84%	59.33%	62.31%
33%	0.27%	13.18%	26.27%	35.97%	43.37%	49.21%	53.93%	57.83%	61.12%	63.93%
35%	0.50%	15.88%	29.36%	38.91%	46.06%	51.62%	56.10%	59.79%	62.88%	65.52%
37%	0.87%	18.71%	32.44%	41.80%	48.67%	53.97%	58.22%	61.70%	64.62%	67.10%
39%	1.39%	21.64%	35.49%	44.62%	51.22%	56.26%	60.28%	63.57%	66.31%	68.65%
41%	2.11%	24.65%	38.51%	47.37%	53.69%	58.49%	62.29%	65.39%	67.97%	70.17%
43%	3.03%	27.71%	41.47%	50.05%	56.09%	60.64%	64.24%	67.16%	69.59%	71.66%
45%	4.17%	30.80%	44.38%	52.65%	58.41%	62.73%	66.12%	68.88%	71.17%	73.11%
47%	5.54%	33.89%	47.22%	55.18%	60.66%	64.75%	67.95%	70.54%	72.70%	74.52%
49%	7.15%	36.96%	49.99%	57.62%	62.83%	66.70%	69.72%	72.15%	74.18%	75.89%
51%	8.98%	40.01%	52.69%	59.98%	64.93%	68.58%	71.42%	73.71%	75.61%	77.21%
53%	11.03%	43.01%	55.30%	62.26%	66.94%	70.39%	73.06%	75.21%	76.99%	78.50%
55%	13.28%	45.96%	57.82%	64.45%	68.88%	72.12%	74.64%	76.65%	78.32%	79.73%
57%	15.70%	48.85%	60.26%	66.56%	70.74%	73.79%	76.15%	78.04%	79.60%	80.92%
59%	18.30%	51.67%	62.61%	68.58%	72.52%	75.39%	77.60%	79.37%	80.83%	82.06%
61%	21.03%	54.40%	64.87%	70.52%	74.23%	76.92%	78.99%	80.65%	82.01%	83.16%
63%	23.87%	57.05%	67.04%	72.37%	75.86%	78.38%	80.32%	81.86%	83.14%	84.21%
65%	26.81%	59.61%	69.11%	74.14%	77.41%	79.77%	81.58%	83.03%	84.22%	85.22%
67%	29.82%	62.08%	71.09%	75.82%	78.89%	81.10%	82.79%	84.14%	85.24%	86.17%
69%	32.88%	64.45%	72.99%	77.43%	80.30%	82.36%	83.94%	85.19%	86.22%	87.09%
71%	35.96%	66.73%	74.79%	78.95%	81.64%	83.56%	85.03%	86.19%	87.15%	87.95%
73%	39.05%	68.90%	76.50%	80.40%	82.91%	84.70%	86.06%	87.14%	88.03%	88.78%
75%	42.12%	70.98%	78.13%	81.77%	84.11%	85.77%	87.04%	88.04%	88.87%	89.56%
77%	45.17%	72.96%	79.67%	83.07%	85.24%	86.79%	87.96%	88.90%	89.66%	90.30%
79%	48.18%	74.84%	81.13%	84.30%	86.31%	87.75%	88.84%	89.70%	90.41%	91.00%
81%	51.13%	76.62%	82.50%	85.45%	87.32%	88.65%	89.66%	90.46%	91.11%	91.66%
83%	54.01%	78.31%	83.80%	86.54%	88.28%	89.50%	90.44%	91.17%	91.78%	92.28%
85%	56.81%	79.91%	85.02%	87.56%	89.17%	90.30%	91.16%	91.85%	92.40%	92.87%
87%	59.53%	81.42%	86.17%	88.52%	90.01%	91.06%	91.85%	92.48%	92.99%	93.42%
89%	62.15%	82.83%	87.25%	89.42%	90.79%	91.76%	92.49%	93.07%	93.54%	93.93%
91%	64.67%	84.16%	88.26%	90.27%	91.53%	92.42%	93.09%	93.62%	94.05%	94.41%
93%	67.09%	85.41%	89.20%	91.05%	92.21%	93.03%	93.65%	94.13%	94.53%	94.86%
95%	69.41%	86.58%	90.08%	91.79%	92.85%	93.60%	94.17%	94.62%	94.98%	95.28%
97%	71.61%	87.68%	90.90%	92.47%	93.45%	94.14%	94.66%	95.06%	95.40%	95.68%
99%	73.70%	88.70%	91.67%	93.11%	94.00%	94.63%	95.11%	95.48%	95.79%	96.04%
101%	75.69%	89.65%	92.38%	93.70%	94.52%	95.09%	95.53%	95.87%	96.15%	96.38%

	Profitability Ratio (% in-the-money)										
Volatility	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
1%	52.76%	57.06%	60.64%	63.66%	66.26%	68.51%	70.48%	72.21%	73.76%	75.14%	76.38%
3%	52.76%	57.06%	60.64%	63.66%	66.26%	68.51%	70.48%	72.21%	73.76%	75.14%	76.38%
5%	52.76%	57.06%	60.64%	63.66%	66.26%	68.51%	70.48%	72.21%	73.76%	75.14%	76.38%
7%	52.78%	57.06%	60.64%	63.66%	66.26%	68.51%	70.48%	72.21%	73.76%	75.14%	76.38%
9%	52.90%	57.11%	60.66%	63.68%	66.26%	68.51%	70.48%	72.21%	73.76%	75.14%	76.38%
11%	53.22%	57.30%	60.77%	63.74%	66.30%	68.53%	70.49%	72.22%	73.76%	75.14%	76.38%
13%	53.78%	57.68%	61.02%	63.91%	66.42%	68.61%	70.54%	72.26%	73.79%	75.16%	76.40%
15%	54.56%	58.26%	61.45%	64.22%	66.65%	68.78%	70.67%	72.36%	73.86%	75.22%	76.44%
17%	55.53%	59.02%	62.04%	64.69%	67.02%	69.08%	70.91%	72.54%	74.01%	75.33%	76.53%
19%	56.65%	59.93%	62.79%	65.30%	67.52%	69.49%	71.25%	72.82%	74.25%	75.53%	76.70%
21%	57.89%	60.96%	63.66%	66.03%	68.14%	70.02%	71.70%	73.21%	74.58%	75.82%	76.95%
23%	59.20%	62.09%	64.63%	66.87%	68.87%	70.65%	72.25%	73.69%	75.00%	76.19%	77.27%
25%	60.58%	63.29%	65.68%	67.80%	69.68%	71.37%	72.88%	74.26%	75.50%	76.64%	77.68%
27%	61.99%	64.55%	66.79%	68.79%	70.56%	72.16%	73.59%	74.89%	76.08%	77.16%	78.15%
29%	63.44%	65.84%	67.95%	69.83%	71.50%	73.00%	74.36%	75.59%	76.72%	77.74%	78.68%
31%	64.89%	67.15%	69.14%	70.90%	72.48%	73.90%	75.18%	76.34%	77.40%	78.38%	79.27%
33%	66.35%	68.47%	70.34%	72.00%	73.49%	74.82%	76.03%	77.13%	78.13%	79.05%	79.90%
35%	67.81%	69.80%	71.56%	73.12%	74.52%	75.77%	76.91%	77.95%	78.89%	79.76%	80.56%
37%	69.25%	71.12%	72.77%	74.24%	75.55%	76.74%	77.81%	78.79%	79.68%	80.50%	81.25%
39%	70.67%	72.43%	73.98%	75.36%	76.60%	77.71%	78.72%	79.64%	80.48%	81.25%	81.96%
41%	72.07%	73.72%	75.18%	76.48%	77.64%	78.68%	79.63%	80.49%	81.28%	82.01%	82.68%
43%	73.44%	74.99%	76.36%	77.58%	78.67%	79.65%	80.54%	81.35%	82.09%	82.78%	83.41%
45%	74.78%	76.24%	77.52%	78.66%	79.69%	80.61%	81.44%	82.21%	82.90%	83.55%	84.14%
47%	76.09%	77.46%	78.66%	79.73%	80.69%	81.55%	82.34%	83.05%	83.71%	84.31%	84.87%
49%	77.36%	78.64%	79.77%	80.77%	81.67%	82.48%	83.22%	83.89%	84.50%	85.06%	85.59%
51%	78.59%	79.79%	80.85%	81.79%	82.63%	83.39%	84.08%	84.70%	85.28%	85.81%	86.30%
53%	79.79%	80.91%	81.90%	82.78%	83.56%	84.27%	84.92%	85.51%	86.04%	86.54%	87.00%
55%	80.94%	81.99%	82.92%	83.74%	84.47%	85.14%	85.74%	86.29%	86.79%	87.25%	87.68%
57%	82.05%	83.03%	83.90%	84.66%	85.35%	85.97%	86.53%	87.05%	87.52%	87.95%	88.35%
59%	83.12%	84.04%	84.84%	85.56%	86.20%	86.78%	87.31%	87.78%	88.22%	88.63%	89.00%
61%	84.15%	85.00%	85.75%	86.42%	87.02%	87.56%	88.05%	88.50%	88.91%	89.28%	89.63%
63%	85.13%	85.93%	86.63%	87.25%	87.81%	88.31%	88.77%	89.18%	89.57%	89.92%	90.24%
65%	86.07%	86.81%	87.47%	88.05%	88.57%	89.03%	89.46%	89.85%	90.20%	90.53%	90.83%
67%	86.97%	87.66%	88.27%	88.81%	89.29%	89.73%	90.12%	90.48%	90.81%	91.11%	91.40%
69%	87.83%	88.47%	89.03%	89.53%	89.98%	90.39%	90.75%	91.09%	91.39%	91.68%	91.94%
71%	88.64%	89.24%	89.76%	90.23%	90.64%	91.02%	91.36%	91.67%	91.95%	92.22%	92.46%
73%	89.42%	89.97%	90.46%	90.89%	91.27%	91.62%	91.94%	92.22%	92.49%	92.73%	92.96%
75%	90.15%	90.67%	91.12%	91.51%	91.87%	92.19%	92.49%	92.75%	93.00%	93.22%	93.43%
77%	90.85%	91.32%	91.74%	92.11%	92.44%	92.74%	93.01%	93.25%	93.48%	93.69%	93.88%
79%	91.51%	91.95%	92.33%	92.67%	92.98%	93.25%	93.50%	93.73%	93.94%	94.13%	94.31%
81%	92.13%	92.53%	92.89%	93.20%	93.49%	93.74%	93.97%	94.18%	94.37%	94.55%	94.71%
83%	92.71%	93.09%	93.42%	93.71%	93.97%	94.20%	94.41%	94.61%	94.78%	94.95%	95.10%
85%	93.26%	93.61%	93.91%	94.18%	94.42%	94.63%	94.83%	95.01%	95.17%	95.32%	95.46%
87%	93.78%	94.10%	94.38%	94.62%	94.84%	95.04%	95.22%	95.39%	95.54%	95.67%	95.80%
89%	94.27%	94.56%	94.81%	95.04%	95.24%	95.42%	95.59%	95.74%	95.88%	96.01%	96.12%
91%	94.72%	94.99%	95.22%	95.43%	95.62%	95.78%	95.94%	96.07%	96.20%	96.32%	96.43%
93%	95.15%	95.39%	95.61%	95.80%	95.97%	96.12%	96.26%	96.39%	96.50%	96.61%	96.71%
95%	95.54%	95.77%	95.97%	96.14%	96.30%	96.44%	96.56%	96.68%	96.79%	96.88%	96.97%
97%	95.91%	96.12%	96.30%	96.46%	96.60%	96.73%	96.85%	96.95%	97.05%	97.14%	97.22%
99%	96.26%	96.45%	96.61%	96.76%	96.89%	97.00%	97.11%	97.21%	97.29%	97.38%	97.45%
101%	96.58%	96.75%	96.90%	97.03%	97.15%	97.26%	97.35%	97.44%	97.52%	97.60%	97.67%

Real Options Analysis Values (30-year maturity at 5% risk-free rate)

	Profitability Ratio (% in-the-money)									
Volatility	-99%	-90%	-80%	-70%	-60%	-50%	-40%	-30%	-20%	-10%
1%	0.00%	0.00%	0.05%	25.62%	44.22%	55.37%	62.81%	68.12%	72.11%	75.21%
3%	0.00%	0.00%	2.62%	25.82%	44.22%	55.37%	62.81%	68.12%	72.11%	75.21%
5%	0.00%	0.02%	6.64%	27.30%	44.34%	55.38%	62.81%	68.12%	72.11%	75.21%
7%	0.00%	0.37%	10.94%	29.75%	45.00%	55.54%	62.85%	68.13%	72.11%	75.21%
9%	0.00%	1.57%	15.31%	32.67%	46.31%	56.06%	63.06%	68.22%	72.15%	75.22%
11%	0.00%	3.71%	19.66%	35.80%	48.08%	57.01%	63.56%	68.49%	72.30%	75.31%
13%	0.00%	6.63%	23.98%	39.03%	50.16%	58.30%	64.37%	69.00%	72.63%	75.53%
15%	0.01%	10.13%	28.22%	42.28%	52.43%	59.86%	65.45%	69.76%	73.16%	75.91%
17%	0.04%	14.03%	32.38%	45.51%	54.80%	61.59%	66.72%	70.71%	73.89%	76.47%
19%	0.18%	18.17%	36.44%	48.70%	57.23%	63.44%	68.15%	71.83%	74.77%	77.18%
21%	0.50%	22.46%	40.40%	51.83%	59.66%	65.36%	69.67%	73.06%	75.78%	78.01%
23%	1.13%	26.82%	44.23%	54.88%	62.08%	67.30%	71.26%	74.37%	76.87%	78.94%
25%	2.15%	31.19%	47.94%	57.84%	64.47%	69.25%	72.88%	75.73%	78.03%	79.94%
27%	3.64%	35.51%	51.51%	60.70%	66.79%	71.18%	74.50%	77.11%	79.23%	80.98%
29%	5.62%	39.75%	54.94%	63.46%	69.06%	73.07%	76.11%	78.51%	80.45%	82.05%
31%	8.09%	43.88%	58.22%	66.10%	71.24%	74.92%	77.70%	79.89%	81.66%	83.13%
33%	11.01%	47.89%	61.36%	68.63%	73.35%	76.71%	79.25%	81.25%	82.87%	84.21%
35%	14.34%	51.74%	64.35%	71.04%	75.36%	78.43%	80.75%	82.57%	84.05%	85.28%
37%	18.02%	55.44%	67.18%	73.34%	77.28%	80.08%	82.19%	83.85%	85.20%	86.32%
39%	21.96%	58.96%	69.87%	75.51%	79.11%	81.66%	83.58%	85.09%	86.31%	87.33%
41%	26.11%	62.31%	72.40%	77.56%	80.84%	83.15%	84.90%	86.27%	87.38%	88.31%
43%	30.39%	65.48%	74.78%	79.49%	82.47%	84.57%	86.15%	87.39%	88.40%	89.24%
45%	34.74%	68.47%	77.01%	81.30%	84.00%	85.90%	87.34%	88.46%	89.37%	90.13%
47%	39.11%	71.28%	79.09%	82.99%	85.44%	87.16%	88.45%	89.47%	90.29%	90.98%
49%	43.44%	73.90%	81.04%	84.57%	86.78%	88.33%	89.50%	90.41%	91.16%	91.77%
51%	47.69%	76.35%	82.84%	86.04%	88.03%	89.43%	90.48%	91.30%	91.97%	92.52%
53%	51.82%	78.63%	84.52%	87.40%	89.19%	90.44%	91.39%	92.12%	92.72%	93.22%
55%	55.80%	80.74%	86.06%	88.65%	90.26%	91.39%	92.23%	92.89%	93.43%	93.87%
57%	59.62%	82.68%	87.49%	89.81%	91.25%	92.26%	93.01%	93.60%	94.08%	94.48%
59%	63.24%	84.47%	88.79%	90.87%	92.16%	93.06%	93.73%	94.26%	94.68%	95.04%
61%	66.67%	86.12%	89.99%	91.85%	92.99%	93.79%	94.39%	94.86%	95.24%	95.55%
63%	69.88%	87.62%	91.08%	92.73%	93.75%	94.46%	95.00%	95.41%	95.75%	96.03%
65%	72.88%	88.99%	92.07%	93.54%	94.45%	95.08%	95.55%	95.91%	96.21%	96.46%
67%	75.67%	90.23%	92.98%	94.28%	95.08%	95.63%	96.05%	96.37%	96.64%	96.86%
69%	78.25%	91.36%	93.79%	94.94%	95.65%	96.14%	96.50%	96.79%	97.02%	97.21%
71%	80.61%	92.38%	94.53%	95.54%	96.16%	96.59%	96.91%	97.16%	97.37%	97.54%
73%	82.78%	93.29%	95.19%	96.08%	96.62%	97.00%	97.28%	97.50%	97.68%	97.83%
75%	84.76%	94.11%	95.78%	96.56%	97.04%	97.37%	97.61%	97.81%	97.96%	98.09%
77%	86.55%	94.85%	96.31%	96.99%	97.41%	97.70%	97.91%	98.08%	98.22%	98.33%
79%	88.17%	95.50%	96.78%	97.37%	97.74%	97.99%	98.18%	98.32%	98.44%	98.54%
81%	89.63%	96.09%	97.20%	97.72%	98.03%	98.25%	98.41%	98.54%	98.64%	98.73%
83%	90.93%	96.60%	97.57%	98.02%	98.29%	98.48%	98.62%	98.73%	98.82%	98.90%
85%	92.10%	97.06%	97.90%	98.28%	98.52%	98.68%	98.81%	98.90%	98.98%	99.04%
87%	93.14%	97.46%	98.18%	98.52%	98.72%	98.86%	98.97%	99.05%	99.12%	99.17%
89%	94.06%	97.81%	98.44%	98.73%	98.90%	99.02%	99.11%	99.18%	99.24%	99.29%
91%	94.87%	98.12%	98.66%	98.91%	99.06%	99.16%	99.24%	99.30%	99.35%	99.39%
93%	95.58%	98.39%	98.85%	99.06%	99.19%	99.28%	99.35%	99.40%	99.44%	99.48%
95%	96.21%	98.63%	99.02%	99.20%	99.31%	99.39%	99.44%	99.49%	99.52%	99.55%
97%	96.76%	98.83%	99.17%	99.32%	99.41%	99.48%	99.53%	99.56%	99.59%	99.62%
99%	97.23%	99.01%	99.29%	99.42%	99.50%	99.56%	99.60%	99.63%	99.65%	99.68%
101%	97.65%	99.16%	99.40%	99.51%	99.58%	99.62%	99.66%	99.69%	99.71%	99.72%

	Profitability Ratio (% in-the-money)										
Volatility	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
1%	77.69%	79.72%	81.41%	82.84%	84.06%	85.12%	86.05%	86.87%	87.60%	88.26%	88.84%
3%	77.69%	79.72%	81.41%	82.84%	84.06%	85.12%	86.05%	86.87%	87.60%	88.26%	88.84%
5%	77.69%	79.72%	81.41%	82.84%	84.06%	85.12%	86.05%	86.87%	87.60%	88.26%	88.84%
7%	77.69%	79.72%	81.41%	82.84%	84.06%	85.12%	86.05%	86.87%	87.60%	88.26%	88.84%
9%	77.69%	79.72%	81.41%	82.84%	84.06%	85.12%	86.05%	86.87%	87.60%	88.26%	88.84%
11%	77.74%	79.75%	81.43%	82.85%	84.07%	85.13%	86.06%	86.88%	87.61%	88.26%	88.84%
13%	77.89%	79.85%	81.49%	82.90%	84.10%	85.15%	86.08%	86.89%	87.62%	88.26%	88.85%
15%	78.17%	80.06%	81.65%	83.02%	84.20%	85.23%	86.13%	86.94%	87.65%	88.29%	88.87%
17%	78.61%	80.40%	81.92%	83.24%	84.37%	85.37%	86.25%	87.03%	87.73%	88.36%	88.93%
19%	79.18%	80.87%	82.31%	83.56%	84.65%	85.60%	86.45%	87.20%	87.88%	88.49%	89.04%
21%	79.88%	81.46%	82.81%	83.99%	85.01%	85.92%	86.73%	87.45%	88.09%	88.68%	89.21%
23%	80.67%	82.14%	83.40%	84.50%	85.47%	86.32%	87.08%	87.77%	88.38%	88.94%	89.44%
25%	81.53%	82.90%	84.07%	85.10%	86.00%	86.80%	87.51%	88.15%	88.73%	89.26%	89.74%
27%	82.45%	83.71%	84.80%	85.75%	86.59%	87.33%	88.00%	88.60%	89.14%	89.63%	90.08%
29%	83.41%	84.57%	85.57%	86.45%	87.22%	87.91%	88.53%	89.09%	89.59%	90.05%	90.47%
31%	84.38%	85.44%	86.36%	87.17%	87.89%	88.53%	89.10%	89.61%	90.08%	90.51%	90.90%
33%	85.35%	86.33%	87.17%	87.92%	88.57%	89.16%	89.69%	90.16%	90.60%	90.99%	91.35%
35%	86.32%	87.21%	87.99%	88.67%	89.27%	89.81%	90.29%	90.73%	91.13%	91.49%	91.83%
37%	87.27%	88.08%	88.79%	89.41%	89.96%	90.46%	90.90%	91.30%	91.67%	92.00%	92.31%
39%	88.19%	88.94%	89.58%	90.15%	90.65%	91.10%	91.51%	91.87%	92.21%	92.52%	92.80%
41%	89.09%	89.76%	90.35%	90.87%	91.33%	91.74%	92.11%	92.44%	92.75%	93.02%	93.28%
43%	89.95%	90.56%	91.10%	91.56%	91.98%	92.35%	92.69%	92.99%	93.27%	93.53%	93.76%
45%	90.77%	91.33%	91.81%	92.23%	92.61%	92.95%	93.25%	93.53%	93.78%	94.01%	94.23%
47%	91.56%	92.06%	92.49%	92.88%	93.22%	93.52%	93.80%	94.05%	94.28%	94.49%	94.68%
49%	92.30%	92.75%	93.14%	93.48%	93.79%	94.07%	94.32%	94.54%	94.75%	94.94%	95.11%
51%	92.99%	93.40%	93.75%	94.06%	94.34%	94.59%	94.81%	95.01%	95.20%	95.37%	95.53%
53%	93.64%	94.01%	94.32%	94.60%	94.85%	95.08%	95.28%	95.46%	95.63%	95.78%	95.92%
55%	94.25%	94.58%	94.86%	95.11%	95.33%	95.53%	95.71%	95.88%	96.03%	96.17%	96.29%
57%	94.82%	95.11%	95.36%	95.59%	95.78%	95.96%	96.12%	96.27%	96.40%	96.53%	96.64%
59%	95.34%	95.60%	95.83%	96.02%	96.20%	96.36%	96.51%	96.64%	96.76%	96.87%	96.97%
61%	95.82%	96.05%	96.25%	96.43%	96.59%	96.73%	96.86%	96.98%	97.08%	97.18%	97.27%
63%	96.27%	96.47%	96.65%	96.81%	96.95%	97.07%	97.19%	97.29%	97.38%	97.47%	97.55%
65%	96.67%	96.85%	97.01%	97.15%	97.27%	97.38%	97.49%	97.58%	97.66%	97.74%	97.81%
67%	97.04%	97.20%	97.34%	97.46%	97.57%	97.67%	97.76%	97.84%	97.91%	97.98%	98.05%
69%	97.38%	97.52%	97.64%	97.75%	97.85%	97.93%	98.01%	98.08%	98.15%	98.21%	98.26%
71%	97.68%	97.81%	97.91%	98.01%	98.09%	98.17%	98.24%	98.30%	98.36%	98.41%	98.46%
73%	97.96%	98.07%	98.16%	98.24%	98.32%	98.38%	98.44%	98.50%	98.55%	98.60%	98.64%
75%	98.20%	98.30%	98.38%	98.45%	98.52%	98.58%	98.63%	98.68%	98.72%	98.76%	98.80%
77%	98.43%	98.51%	98.58%	98.64%	98.70%	98.75%	98.80%	98.84%	98.88%	98.91%	98.95%
79%	98.62%	98.70%	98.76%	98.81%	98.86%	98.91%	98.95%	98.98%	99.02%	99.05%	99.08%
81%	98.80%	98.86%	98.92%	98.97%	99.01%	99.05%	99.08%	99.11%	99.14%	99.17%	99.19%
83%	98.96%	99.01%	99.06%	99.10%	99.14%	99.17%	99.20%	99.23%	99.25%	99.28%	99.30%
85%	99.10%	99.14%	99.18%	99.22%	99.25%	99.28%	99.31%	99.33%	99.35%	99.37%	99.39%
87%	99.22%	99.26%	99.29%	99.33%	99.35%	99.38%	99.40%	99.42%	99.44%	99.46%	99.47%
89%	99.33%	99.36%	99.39%	99.42%	99.44%	99.46%	99.48%	99.50%	99.52%	99.53%	99.54%
91%	99.42%	99.45%	99.48%	99.50%	99.52%	99.54%	99.55%	99.57%	99.58%	99.60%	99.61%
93%	99.50%	99.53%	99.55%	99.57%	99.59%	99.60%	99.62%	99.63%	99.64%	99.65%	99.66%
95%	99.58%	99.60%	99.62%	99.63%	99.65%	99.66%	99.67%	99.68%	99.69%	99.70%	99.71%
97%	99.64%	99.66%	99.67%	99.69%	99.70%	99.71%	99.72%	99.73%	99.74%	99.75%	99.75%
99%	99.69%	99.71%	99.72%	99.73%	99.74%	99.75%	99.76%	99.77%	99.78%	99.78%	99.79%
101%	99.74%	99.75%	99.76%	99.77%	99.78%	99.79%	99.80%	99.81%	99.81%	99.82%	99.82%

Answers to End of Chapter Questions and Exercises

Chapter 1 Moving Beyond Uncertainty

1. Risk is important in decision making as it provides an added element of insight into the project being evaluated. Projects with higher returns usually carry with them higher risks, and neglecting the element of risk means that the decision maker may unnecessarily select the riskiest projects.
2. Bang for the buck implies selecting the best project or combination of projects that yields the highest returns subject to the minimum amount of risk. That is, given some set of risk, what is the best project or combination of projects that provide the best returns? Conversely, it also answers what the minimum level of risk is, subject to some prespecified level of return. This concept is the Markowitz efficient frontier in portfolio optimization discussed later in the book.
3. Uncertainty implies an event's outcome in which no one knows for sure what may occur. Uncertainties can range from the fluctuation in the stock market to the occurrences of sunspots. In contrast, uncertainties that affect the outcome of a project or asset's value directly or indirectly are termed risks.

Chapter 2 From Risk to Riches

1. The efficient frontier was first introduced by Nobel laureate Harry Markowitz, and it captures the concept of bang for the buck, where projects or assets are first grouped into portfolios. Then, the combinations of projects or assets that provide the highest returns subject to the varying degrees of risk are calculated. The best and most efficient combinations of projects or assets are graphically represented and termed the *efficient frontier*.
2. Inferential statistics refers to the branch of statistics that performs statistical analysis on smaller-size samples to infer the true nature of the population. The steps undertaken include designing the experiment, collecting the data, analyzing the data, estimating or predicting alternative

conditions, testing of the hypothesis, testing of goodness-of-fit, and making decisions based on the results.

3. Standard deviation measures the average deviation of each data point from the mean, which implies that both upside and downside deviations are captured in a standard deviation calculation. In contrast, only the downside deviations are captured in the semi-standard deviation measure. The semi-standard deviation when used as a measure of risk is more appropriate if only downside occurrences are deemed as risky.
4. Holding everything else constant, projects with negative skew are preferred as the higher probability of occurrences are weighted more on the higher returns.
5. The answer depends on the type of project. For instance, for financial assets such as stocks, clearly a lower kurtosis stock implies a lower probability of occurrence in the extreme areas, or that catastrophic losses are less likely to occur. However, the disadvantage is that the probability of an extreme upside is also lessened.
6. Value at Risk (VaR) measures the worst-case outcome for a particular holding period with respect to a given probability. For instance, the worst-case 5 percent probability VaR of a particular project is $1 million for a 10-year economic life with a 90 percent statistical confidence. Compare that to a simplistic worst-case scenario, which in most cases are single-point estimates, for example, the worst-case scenario for the project is a $10,000 loss. Worst-case scenarios can be added to probabilistic results as in the VaR approach but are usually single-point estimates (usually just a management assumption or guesstimate).

Chapter 3 A Guide to Model-Building Etiquette

For the answers to Chapter 3's Exercises, refer to the enclosed CD-ROM. The files are located in the folder: *Answers to End of Chapter Questions and Exercises*.

Chapter 4 On the Shores of Monaco

1. Parametric simulation is an approach that requires distributional parameters to be first assigned before it can begin. For instance, a Monte Carlo simulation of 1,000 trials using input assumptions in a normal distribution with an average of 10 and standard deviation of 2 is a parametric simulation. In contrast, nonparametric simulation uses historical or comparable data to run the simulation, where specific distributional assumptions (i.e., size and shape of the distribution, type of distribution and its related inputs such as average or standard deviation, and so forth) are not required. Nonparametric simulation is used when the data is "left alone to tell the story."

2. The term "stochastic" means the opposite of "deterministic." Stochastic variables are characterized by their randomness, for example, a stock's price movement over time. A stochastic process is a mathematical relationship that captures this random characteristic over time. The most common stochastic process is the Brownian Motion or random walk used to simulate stock prices.
3. The *RAND()* function in Excel creates a random number from the uniform distribution between 0 and 1. Hitting the F9 key repeatedly will generate additional random numbers from the same distribution.
4. The *NORMSINV()* function in Excel calculates the inverse of the standard cumulative normal distribution with a mean of zero and a standard deviation of one.
5. When used in conjunction, the function *NORMSINV(RAND())* simulates a standard normal distribution random variable.

Chapter 5 Test Driving Risk Simulator

1. Starting a new profile is like starting a new file in Excel, but a profile is part of the Excel file and holds all the information on the simulation parameters; that is, you can perform scenario analysis on simulation by creating multiple similar profiles and changing each profile's distributional assumptions and parameters and see what the resulting differences are.
2. Pearson's product moment correlation coefficient is a linear parametric correlation where the two variables being correlated are assumed to be linearly related and the underlying assumption is that the correlation's distribution is normal. Spearman's rank-based correlation is a nonparametric correlation that can account for nonlinearities between variables and is hence more robust and better suited for use in simulation where different distributions can be correlated to one another due to their nonparametric properties that do not rely on the normal assumption.
3. More simulation trials are required to obtain a lower error level, a higher precision level, and a narrower confidence interval.
4. Error and precision are related but at the same time, they are not the same thing. Error relates to how far off a particular value is, that is, its forecast interval. For example, the mean is 10 with an error of 1, which means that the forecast interval is between 9 and 11. However, precision indicates the level of confidence of this forecast interval. For example, this error has a 90 percent precision, which means that 90 percent of the time, the error will be between 9 and 11.
5. Yes. Even using rough rules of thumb such as ±0.25 (low correlation), ±0.50 (moderate correlation), and ±0.75 (strong correlation) when in

fact there are correlations among the variables although their exact values are unknown, will provide better estimates than not applying these correlations.

Chapter 6 Pandora's Toolbox

1. Tornado and spider charts are used to obtain the static sensitivities of a variable to its precedents by perturbing each of the precedent variables one at a time at a prespecified range. They are typically applied before a simulation is run and no simulation assumptions are required in the analysis. In contrast, sensitivity analysis is applied after a simulation run and requires both assumptions and forecasts. The assumptions are applied in a dynamic environment (with the relevant correlations and truncations) and the sensitivities of the forecast to each of the assumptions are then computed.
2. Some of the distributions are fairly closely related to one another (for instance, the Poisson and binomial distributions become normally distributed when their rates and number of trials increase) and it will be no surprise that some other distribution may be a better fit. In addition, distributions like the beta are highly flexible and can assume multiple shapes and forms, and hence, can be used to fit multiple distributions and data sets.
3. A hypothesis test is used to test if a certain value or parameter is similar to or different from another hypothesized value—for example, whether two means from two different distributions are statistically similar or different.
4. Bootstrap simulation is used to obtain a forecast statistic's confidence interval and hence can be used to determine a statistic's precision and error level.
5. The square of the nonlinear rank correlation coefficient is an approximation of the percent variation in a sensitivity analysis.

Chapter 8 Tomorrow's Forecast Today

1. Time-series forecasting can be used to incorporate linear trends and seasonality in the forecasts while nonlinear extrapolation can only incorporate a nonlinear trend in its forecast. The former cannot include a nonlinear trend while the latter cannot have a seasonality component in its forecasts.
2. All forecasting methods require data except for stochastic process forecasts, which do not require any historical or comparable data, albeit the existence of data can be exploited by using these data to compute the relevant growth rate, volatility, reversion rate, jump rates, and so forth, used in generating these stochastic processes.

3. A Delphi survey method can be applied and the results of the survey can be used to generate a custom distribution. Simulation can, hence, be applied on this custom distribution.
4. Go through and replicate the examples in the chapter.
5. This statement is true. Seasonality is, in most cases, easy to forecast, but cyclicality is more difficult if not impossible to forecast. Examples of seasonality effects include the sales levels of ski passes (peaks during winter and troughs during summer and, hence, are fairly easy to predict year after year) versus cyclicality effects like the business cycle or stock price cycles (extremely hard to predict as the timing, frequency, and magnitude of peaks and troughs are highly unpredictable).

Chapter 9 Using the Past to Predict the Future

1. a. Time-series analysis
   The application of forecasting methodology on data that depends on time.
   b. Ordinary least squares
   A type of regression analysis that minimizes the sum of the square of errors.
   c. Regression analysis
   The estimation of the best-fitting line through a series of historical data used to predict a statistical relationship or to forecast the future based on this relationship.
   d. Heteroskedasticity
   The variance of the errors of a regression analysis is unstable over time.
   e. Autocorrelation
   The historical data of a variable depends on or is correlated to itself over time.
   f. Multicollinearity
   The independent variables are highly correlated to each other or there exists an exact linear relationship between the independent variables.
   g. ARIMA
   Autoregressive Integrated Moving Average—a type of forecasting methodology.
2. The R-squared or coefficient of determination is used on bivariate regressions, whereas the adjusted R-squared is used on multivariate regressions. The latter penalizes the excessive use of independent variables through a degree of freedom correction, making it a more conservative measure useful in multivariate regressions.

3. a. Heteroskedasticity
   In the event of heteroskedasticity, the estimated R-squared is fairly low and the regression equation is both insufficient and incomplete, leading to potentially large estimation errors.
   b. Autocorrelation
   If autocorrelated dependent variable values exist, the estimates of the slope and intercept will be unbiased, but the estimates of their variances will not be reliable and hence the validity of certain statistical goodness-of-fit tests will be flawed.
   c. Multicollinearity
   In perfect multicollinearity, the regression equation cannot be estimated at all. In near-perfect collinearity, the estimated regression equation will be inefficient and inaccurate. The corresponding R-squared is inflated and the t-statistics are lower than actual.
4. Nonlinear independent variables can be transformed into linear variables by taking the logarithm, square (or higher powers), square root, or multiplicative combinations of the independent variables. A new regression is then run based on these newly transformed variables.

Chapter 10 The Search for the Optimal Decision

1. Deterministic optimization means that the input variables are single-point deterministic values, whereas optimization under uncertainty means that the input variables are uncertain and simulated while the optimization process is occurring.
2. a. Objective
   An objective is the forecast output value that is to be maximized or minimized in an optimization (e.g., profits).
   b. Constraint
   A constraint is a restriction that is observed in an optimization (e.g., budget constraint).
   c. Decision variable
   The variables that can be changed based on management decisions such that the objective is achieved. These variables are usually subject to the constraints in the model.
3. Some problems arising from a graphical linear programming approach include nonlinear constraints, unbounded solutions, no feasible solutions, multiple solutions, and too many constraints. These problems cannot be easily solved graphically.
4. The graphical approach is simple to implement but may sometimes be too tedious if too many constraints or nonlinear constraints exist. Optimization can also be solved mathematically by taking first and second

derivatives but is more difficult to do. Excel's Solver add-in can be used to systematically search by brute force through a series of input combinations to find the optimal solution, but the results may be local minimums or local maximums, providing incorrect answers. Risk Simulator also can be used to solve an optimization problem under uncertainty when the input assumptions are unknown and simulated.

Chapter 11 Optimization Under Certainty

1. Discrete decision variables are typically integers such as 0, 1, 2, 3, and so forth, whereas continuous variables can vary between any two values (e.g., between 0 and 1, we can have an infinite number of values such as 0.113354, 0.00012546, and so forth).

Chapter 12 What Is So Real About Real Options, and Why Are They Optional?

1. Real options analysis is an integrated risk analysis process that is used to hedge risks and to take advantage of upside uncertainties, and is used for strategic decision analysis.
2. Real options can be solved using closed-form models, simulation approaches, binomial and multinomial lattices, as well as other more advanced numerical approaches such as variance reduction.
3. The method must be valid, accurate, replicable, tractable, robust, explainable, and most importantly, flexible enough to handle various inputs and able to mirror real-life conditions.
4. A model must exist or can be built; there must exist uncertainties and risks in the decision; these uncertainties and risks must affect the outcomes and hence the decisions in the project; there must be strategic flexibility or options in the project; and the decision makers or senior management must be credible enough to execute the options when they become optimal to do so.
5. The risks and uncertainties are not hedged or taken advantage of; that is, simulation can be used to forecast, predict, and quantify risks, but only real options analysis can be applied to hedge these risks or to take advantage of the upside.

Chapter 15 The Warning Signs

1. "Negligent entrustment" simply means that management takes the results from some fancy analytics generated by an analyst as is, without any due diligence performed on them. This situation usually occurs because management does not understand the approach used or know the relevant questions to ask.

2. Some general types of errors encountered when creating a model include model errors, assumption and input errors, analytical errors, user errors, and interpretation errors.
3. If truncation is not applied when it should be, then the resulting forecast distribution will be too wide and the errors of estimations too large. Therefore, truncation is important as it provides results that are more accurate with lower errors.
4. A critical success factor is an input variable that has significant impact on the output result. By itself, the input variable is also highly uncertain and should be simulated.
5. A skewness of 0 and a kurtosis of 3 or excess kurtosis of 0 are considered normal-looking statistics.
6. Structural breaks occur when the underlying variable undergoes certain economic, business, or financial shifts (e.g., merger or divestiture). Specification errors occur when the underlying variable follows some nonlinearities (e.g., growth curves, exponential, or cyclical curves) but the regression is estimated based on a strict linear model.

About the CD-ROM

INTRODUCTION

The CD-ROM accompanying this book contains trial versions of the Risk Simulator and Real Options Super Lattice Solver (SLS) software, and several Getting Started videos showing how to use the software. The ReadMe file at the root of the CD-ROM provides some quick introductions to the system requirements and installation instructions. Refer to the book for more details on using the software and interpreting the results. In addition, you can visit the software developer's web site at www.realoptionsvaluation.com for more software details, to obtain free models and information about training and certifications, and to view the frequently asked questions about the software.

SYSTEM REQUIREMENTS

- IBM PC or compatible computer with Pentium III or higher processor.
- 128 MB RAM (256 MB recommended) and 30 MB hard-disk space.
- CD-ROM drive, SVGA monitor with 256 colors.
- Excel XP or 2003.
- Windows 2000, XP (preferred), or higher.
- Microsoft .NET Framework 1.1.
- Administrative privileges to install the software (not applicable on home computers).
- An Internet connection.

USING THE CD WITH WINDOWS

There is an automated setup program available in the CD-ROM for the Real Options Super Lattice Solver and the Risk Simulator software. You must first be connected to the Internet to install these software. To run the setup program, do the following:

1. Insert the enclosed CD into the CD-ROM drive of your computer.

2. The setup program should come up automatically. If it does not, open Windows Explorer and double click the CDAutorun.exe file to launch the interface.
3. Click on Install Real Options SLS, and click Run. Follow the on-screen instructions.
4. When prompted, enter the following user name and license key for a 30-day trial of the SLS software:

   Name: 30 Day License   Key: 513C-27D2-DC6B-9666

5. Then go back and click on Install Risk Simulator, and click Run. Follow the on-screen instructions. Make sure you have the specified system requirements before installing the software.

To obtain a permanent license or an extended academic trial (a special offer for professors and students), contact Admin@RealOptionsValuation.com for details.

CONTENT

The enclosed CD-ROM contains 30-day trial versions of the Real Options Super Lattice Solver and the Risk Simulator software, and Getting Started videos.

This CD contains an installer for the Real Options Super Lattice Solver software and Risk Simulator software. To install these software, insert the CD-ROM and the installer will automatically appear. If it does not, browse this CD and double click on CDAutorun.exe. You must first be connected to the Internet before installation can occur, as the installer will be downloading the latest setup files for these software from the Real Options Valuation, Inc., web site (www.realoptionsvaluation.com).

Finally, Risk Simulator requires Microsoft .NET Framework 1.1 installed to function. Most new computers come with .NET Framework preinstalled. For older machines, you may have to install it manually. The DOT NET Framework folder in the CD has a file called dotnetfx.exe.

Install this file if you do not have .NET Framework 1.1 preinstalled. If you do not know if you have .NET Framework on your computer, you can install this file just to make sure.

Trial, demo, or *evaluation* versions are usually limited either by time or functionality (such as being unable to save projects). Some trial versions are very sensitive to system date changes. If you alter your computer's date, the programs will "time out" and no longer be functional.

CUSTOMER CARE

If you have trouble with the CD-ROM, call the Wiley Product Technical Support phone number at (800) 762-2974. Outside the United States, call 1(317) 572-3994. You can also contact Wiley Product Technical Support at http://support.wiley.com. John Wiley & Sons will provide technical support only for installation and other general quality control items. For technical support on the applications themselves, consult the program's vendor or author.

To place additional orders or to request information about other Wiley products, call (877) 762-2974.

If you have questions about obtaining the fully functional software, contact Dr. Johnathan Mun at JohnathanMun@cs.com.

Index

For more information about the CD-ROM, see the
About the CD-ROM section on page 588.

CUSTOMER NOTE:

IF THIS BOOK IS ACCOMPANIED BY SOFTWARE, PLEASE READ THE FOLLOWING BEFORE OPENING THE PACKAGE.

WILEY