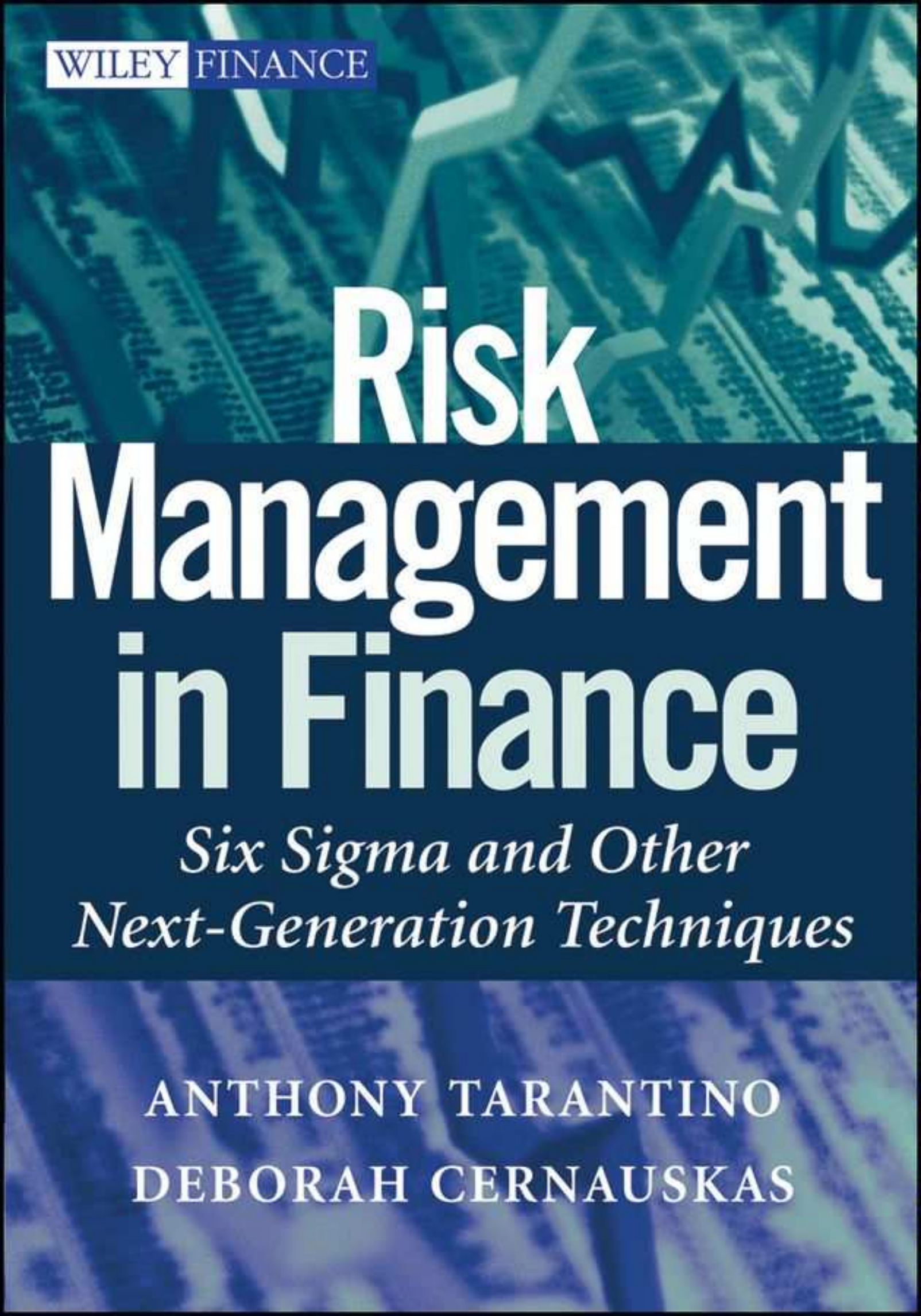

WILEY FINANCE
Risk Management in Finance
Six Sigma and Other Next-Generation Techniques
ANTHONY TARANTINO
DEBORAH CERNAUSKAS

Risk Management in Finance

Six Sigma and Other Next-Generation Techniques

ANTHONY TARANTINO
DEBORAH CERNAUSKAS

WILEY
John Wiley & Sons, Inc.

Risk Management in Finance

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, please visit our Web site at www.WileyFinance.com.

Risk Management in Finance

Six Sigma and Other Next-Generation Techniques

ANTHONY TARANTINO
DEBORAH CERNAUSKAS

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:

Tarantino, Anthony, 1949–
Risk management in finance : six sigma and other next generation techniques / Anthony Tarantino, Deb Cernauskas.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-470-41346-3 (cloth)
1. Financial risk management. I. Cernauskas, Deb, 1956– II. Title.
HG173.T346 2009
658.15′5–dc22
2008052035

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

To Winkey, Peapod, and SanSan
—A.T.

To Mom for her continued support
—D.C.

Contents

Preface

According to the Book of Genesis, God decided to destroy the world in a great flood because of mankind's sinful and wicked ways. But God knew Noah was a righteous man and decided to spare him and his family. He instructed Noah to build an ark, a very large vessel of no economic or recreational value, to hold Noah's family and representatives from the animal kingdom. While there was no business case or quantitative or qualitative risk model to justify this endeavor, Noah decided to mitigate his risk and build the ark. We can imagine that conventional wisdom of the time condemned Noah for such a foolish waste of time and money and that community and media reaction would have been very negative as well.

Noah's risk mitigation proved to be quite timely as conventional wisdom and traditional risk management failed in a catastrophic manner. Noah survived the great flood and began rebuilding civilization after the waters of the great flood receded.

Some time later, Toyota, a Japanese car manufacturer, decided to build a hybrid car to mitigate the risk of rising fuel prices and need to curtail greenhouse gases. As with Noah, there was no valid business case or accepted risk model to justify such a foolish waste of time and money. Conventional wisdom of the time was that large gas-guzzling vehicles were the safe choice. They were all the rage and generated very high returns. Fuel-efficient cars were much less profitable and lacked the status and prestige of larger and more muscular vehicles. As with Noah, we can imagine industry leaders making fun of such a wimpy car that would appeal only to a small number of tree-hugging environmentalists on the American West Coast.

Again, conventional wisdom and traditional risk management failed in a catastrophic manner. The energy crisis and push for green energy made the little hybrid car a huge success and helped propel Toyota into a leadership position as the most profitable and best-capitalized manufacturer in the industry. Conversely, their American competitors are now on the verge of bankruptcy and capitalized below their World War II levels.

A few years ago, Wells Fargo decided that the risk inherent in the subprime mortgage market was unacceptable, and minimized their exposure. Again, the conventional wisdom and accepted quantitative and qualitative risk models argued against their conservatism. Profit margins for subprime mortgages, mortgage-backed securities, and credit default swaps were much higher than the more traditional vehicles and instruments offered by banks. Government regulators, rating agencies, and business media all promoted the subprime market, either directly or indirectly. This created shareholder pressures to jump into this very lucrative market. As with Noah and Toyota, media and public reaction was negative to Wells Fargo's conservative approach to risk mitigation. As with Noah and Toyota, we can imagine industry leaders making fun of a bank with a stagecoach as a corporate symbol—too sentimental and old fashioned to grasp the huge profit potentials in subprime.

Once again, conventional wisdom and traditional risk management failed in a catastrophic manner. Wells Fargo not only survived the global crisis, but substantially expanded its market position. Those who embraced subprime and its related products have been forced out of business or critically wounded. Their subprime activities have brought about the greatest financial crisis since the Great Depression of the 1930s. Unlike Toyota, their failures in risk management negatively impacted the global economy.

Our three parables demonstrate that risk management is never as easy or predicable as conventional wisdom would lead one to believe. Each catastrophic failure in risk management brings greater focus on the need for more innovative and effective techniques for risk management. Unfortunately, memories are short, and new opportunities continue to arise and overwhelm sound risk management.

Financial risk management is especially challenging. Today's financial products and markets are too complex and opaque for the regulatory structures, audit practices, rating agencies, and risk management in place to oversee and control them. Business and accounting schools struggle to keep pace in their curricula with such a dynamic market. Government regulatory structures, designed in the Great Depression, were particularly ineffective in grasping the danger that very complex and highly leveraged financial products presented not just to the banking industry but to all of society. Rating agencies never predicted the collapse of firms, even when the evidence became obvious. Auditors who focused on tactical internal controls regulated under the Sarbanes-Oxley Act failed to grasp the systemic risks that financial services faced.

Noah, Toyota, and Wells Fargo share some important characteristics. All three defied conventional wisdom and public pressure to pursue major opportunities—for an immoral lifestyle during Noah's time, for big gas-guzzling cars during Toyota's time, and for subprime mortgages during Wells Fargo's time. All three did the morally and ethically correct thing: Noah led a righteous life, Toyota helped to fight greenhouse gases, and Wells Fargo declined to market loans that eventually cost millions of borrowers their homes. Each also utilized risk management in a unique manner as compared to their peers that provided a strategic competitive advantage. Staying alive in the case of Noah and prospering economically in the case of Toyota and Wells Fargo.

Financial risk management applies a systematic and logical approach to uncertainties in operations, reputations, credit, and markets. Without risk management, an organization would simply rely on luck to avoid disasters. Financial risk management as a discipline has progressed since the pivotal year of 1921, when Frank Knight published his *Risk, Uncertainty and Profit* and John Maynard Keynes published his *A Treatise on Probability*. Knight pioneered the notion that uncertainty, which cannot be measured, is different from risk, which is measurable. Keynes pioneered the mathematical and philosophical foundations to risk management. Keynes argued for a greater reliance on perception and judgment when considering probabilities and warned of an overreliance on numbers.[1,2,3]

In 1956, Russell Gallagher published his "Risk Management: A New Phase of Cost Control," in the *Harvard Business Review*. As an insurance executive, he argued that a professional insurance manager should also be a risk manager. Because of the nature of its business, the insurance industry was the first to embrace professional risk management with its concern for avoiding unaffordable potential losses. This

leadership continued into the 1960s and 1970s when the Insurance Institute of America developed a certification examination and designation process for an "Associate in Risk Management," and when insurance executives formed the Geneva Association, which advocated the links among risk management, insurance, and economics.

In the 1980s, new risk societies were created to promote risk management—the Society for Risk Analysis in Washington, and the Institute for Risk Management in London. Their efforts have made the concepts of risk assessments and risk management well understood in business and government circles.

In the 1990s, the United Kingdom's Cadbury and Turnbull committees issued reports advocating that corporate boards take responsibility for setting risk management policies, for assuring that the organization understands all its risks, and for accepting oversight for the entire process. It was also in the 1990s that the title chief risk officer (CRO) is first used by GE Capital to describe a manager who is responsible for the totality of risk exposure to an organization. Chief risk officers and risk managers are now commonplace in the financial services industry and spreading into other industries.

The global financial crisis of 2007–2008 begs the question, with all the progress in risk management, why were the world's leading financial services firms, their regulators, their auditors, and their rating agencies so wrong in their assessment of the inherent risks in the subprime mortgage market? These organizations possessed the most sophisticated risk management processes and technologies in the hands of the best-educated and trained risk managers. We believe that part of the reason was that they have not deployed the next-generation techniques we provide here. These techniques could have helped to reduce the pain of the current crisis, and provide risk, business, and IT managers with tools and solutions to substantially improve their risk mitigation. There have always been leaders such as Noah, Toyota, and Wells Fargo, who innovated in their risk management. Hopefully, our suggestions and recommendations will help your organization become innovators as well. As the current global crisis and our three parables demonstrate, this can mean much more than providing a strategic advantage. It can mean the survival of an organization.

The problem with risk management can be summarized in the teachings of the legendary Samurai master swordsman Miyamoto Musashi, in his *Book of the Five Rings*. Musashi won over 30 duels and warned to never take too hard a focus on the point of your opponent's sword. While this would seem to be the obvious point of attack and the greatest risk, the attack always comes from some other point. Therefore, a swordsman must maintain a soft focus to look at the entire field of view. Risk is like this. The biggest threats never come from the most visible point of attack. This was true for Noah's neighbors, Toyota's fellow carmakers, and Wells Fargo's fellow banks.

This is my third book for John Wiley & Sons targeting governance, risk, and compliance. The three books are written as a series and designed to complement each other:

- *The Manager's Guide to Compliance* focuses on the basics of compliance with overviews of best practice frameworks, governance, and audit standards.
- *Governance, Risk, and Compliance Handbook* focuses on the largest economies, regions, and industries in the world as to their corporate, environmental, and

information technology (IT) governance, regulatory compliance, and operational risk management.
- *Risk Management in Finance: Six Sigma and Other Next-Generation Techniques* focuses exclusively on next-generation techniques to improve operational risk management.

Your comments and suggestions are always welcome. E-mail me at agtarantino@hotmail.com, or at my web site, AnthonyTarantino.com.

NOTES

1. Wikipedia, "Frank Knight," http://en.wikipedia.org/wiki/Frank_Knight (accessed November 2008).
2. Wikipedia, "John Maynard Keynes," http://en.wikipedia.org/wiki/John_Maynard_Keynes (accessed November 2009).
3. See "A Short History of Risk Management: 1900 to 2002," www.mccombs.utexas.edu/dept/irom/bba/risk/rmi/arnold/downloads/Hist_of_RM_2002.pdf.

Acknowledgments

We wish to acknowledge the tremendous contributions of our collaborators to this text. Their efforts have produced leading-edge thought leadership based on innovative problem solving and research. They come from a wide variety of backgrounds but share our passion for advancing risk management and corporate governance.

We also wish to acknowledge the support and encouragement of our Wiley colleagues and friends: Tim Burgard, our senior editor; Helen Cho, our editorial coordinator; and Stacey Rivera, our development editor.

About the Contributors

Brian Barnier is a leader at IBM on IT risk and return performance. In this role, he helps the IBM CIO organization and external clients improve alignment between business strategy and model, IT goals and objectives, and business outcomes through a more risk-aware approach to IT investment priorities. He has been an adjunct professor in operations management and finance, serves on several industry standards and practices bodies, teaches continuing professional education sessions, and writes. He coholds the copyright on the Value Added Diamond business performance model and led teams to seven U.S. patents. For more information, you can contact him at bbarnier@us.ibm.com.

Ying Chen, Ph.D., is a master inventor, research staff member, and manager in IBM Almaden Services Research. Ying received her Ph.D. from the Computer Science Department at the University of Illinois at Urbana-Champaign in 1998. She has over 10 years of industry experience in an established IBM research center and a storage start-up company. Her research interests are primarily in information analytics and service-oriented architecture. She also has extensive backgrounds in storage systems, parallel and distributed computing, databases, performance evaluation, and modeling. Ying is currently leading a global research team to develop and deliver successful information analytics solutions and platforms, such as Business Insights Workbench (BIW), which resulted in multimillion-dollar business impact in IBM.

Jill Eicher is a managing director of Adaptive Alpha LLC, a Chicago-based innovator in quantitative analytics arming institutional investors with tools to uncover and profit from dynamic risk opportunities. A seasoned chief operating officer, Ms. Eicher's 25-year career in the investment industry has focused on managing investment businesses competitively by optimizing risk/reward decision making and execution. Her patented risk methodology serves as the foundation of the company's research-and-development platform.

Pedro Fabiano is currently senior vice president at MDB International in Alexandria, Virginia. He is responsible for fraud investigations and prevention, fraud risk consulting, compliance, and related training activities, particularly as they pertain to U.S. companies with interests in Latin America. Mr. Fabiano has more than 15 years of international experience in overseeing governance, compliance, and risk-related matters for U.S. entities in Latin America. Mr. Fabiano is a Regent Emeritus and Fellow of the Association of Certified Fraud Examiners (ACFE). He has authored the "International Bribery" course published by the ACFE, which is used to train professionals around the world.

Allan D. Grody has had hands-on experience in multiple sectors of the financial industry and has been consulting domestically and internationally on issues related to financial institutions' global strategies, restructuring and acquisition needs, capital

and contract market structures, information systems, communications networking, and risk management methods and systems.

As an entrepreneur, he founded his current firm, Financial InterGroup, over two decades ago. Financial InterGroup Advisors is a strategy and acquisition consultancy, advising financial enterprises and their technology suppliers. Financial InterGroup Holdings is a financial industry development company that created six start-ups and formed joint ventures with exchanges and clearinghouses and global technology companies.

He is the author or coauthor of many papers and articles on risk management. He has represented firms in regulatory and trading matters before the Securities and Exchange Commission (SEC); has counseled with trade associations, exchanges, and technology companies; and was an expert witness in a number of financial industry trading patent cases and investment company shareholder suits. He was a member of the board of directors of the technology committee of the Futures Industry Association; an executive committee member of the Emerging Business Council of the Information Industry Association; an executive board member of the Vietnamese Capital Markets Committee and, for nearly a decade, an advisory board member to the London Stock Exchange's Computers in the City Conference. He is currently an editorial board member of the *Journal of Risk Management in Financial Institutions*.

Praveen Gupta, a management consultant, has authored several books, including *Business Innovation in the 21st Century*, *Stat Free Six Sigma*, *The Six Sigma Performance Handbook,* and *Service Scorecard*. He is the editor-in-chief of the *International Journal of Innovation Science,* and writes a monthly column, "Manufacturing Excellence," in *Quality Magazine*. He frequently speaks at conferences internationally. Praveen has been recognized as a thought leader in areas of excellence and innovation and has developed the Six Sigma Scorecard, the 4P model of excellence, Breakthrough innovation, and Stat Free Six Sigma methods that have been translated around the world. Praveen, the founding president of Accelper Consulting (www.accelper.com), has worked at Motorola and AT&T Bell Laboratories, and consulted with about 100 small to large-sized companies including CNA and Abbott Labs. Praveen has taught operations management at DePaul University and business innovation at the Illinois Institute of Technology, Chicago. He has conducted seminars worldwide for over 20 years. Accelper Consulting provides training and consulting services in the area of innovation, Six Sigma, and business performance for achieving sustained profitable growth.

Jeffrey T. Hare is a respected expert on internal controls and security for ERP systems. His background includes public accounting (including Big 4 experience), industry, and Oracle Applications consulting. Jeff has been working in the Oracle Applications space since 1998. His focus is solely on the development of internal controls and security best practices for companies running Oracle Applications. Jeff is a certified public accountant (CPA), a certified information systems auditor (CISA), and a certified internal auditor (CIA). Jeff has worked in various countries, including Australia, Canada, Mexico, Brazil, the United Kingdom, and Germany. Jeff is a graduate of Arizona State University and lives in northern Colorado with his wife and three daughters. You can reach him at jhare@erpseminars.com or (602) 769-9049.

Peter J. Hughes is a chartered accountant; a former country/area executive with JPMorgan Chase; managing director/cofounder of ARC Best Practices Limited, established in 2002; and a principal of the Financial InterGroup Companies. Mr. Hughes accumulated vast experience and knowledge of banks and banking through his 26-year career with JPMorgan Chase, which he has since put to very good use in his career as an independent consultant and adviser. At JPMorgan Chase he was the Central European deputy regional audit manager in their Frankfurt office, South American regional audit manager in their Rio de Janeiro office, country operations executive (Brazil), country senior financial officer (Brazil), country chief administrative officer (Germany), country head of treasury and trading (Germany), head of Europe finance shared services and head of risk management–global shared technology and operations. He was a member of the board of Banco Chase Manhattan SA, Brazil; member of the board (Aufsichtsrat) of Chase Leasing & Co. KG, Germany; and the Chase Manhattan Bank NA, Frankfurt branch manager.

As an independent consultant, Mr. Hughes has advised a number of leading banks, global IT companies and consulting firms, trade associations, and banking institutes. While at JPMorgan Chase, Mr. Hughes pioneered the concept of using business process information and transaction data as a basis for measuring exposure to cross-enterprise risks and the effectiveness of risk mitigation systems. He subsequently collaborated with Allan D. Grody in research and advisory projects involving some of the globe's leading IT and consulting firms, with particular emphasis on risk measurement and management systems and Basel II.

Mr. Hughes is the author/coauthor of a number of academic papers, including "The Direct Measurement of Exposure and Risk in Bank Operations" published in the *Journal of Risk Management in Financial Institutions* and, with Allan D. Grody and Dr. Robert M. Mark, "Operational Risk, Data Management, and Economic Capital" published in the *Journal of Financial Transformation*, Cass-Capco Institute Paper Series on Risk. He was also featured in the industry best-selling book *Operational Risk—Practical Approaches to Implementation*, published by Incisive Media. For many years he represented JPMorgan Chase on the British Bankers' Association's Op Risk Advisory Panel. He is a regular speaker at conferences and presents training courses and workshops on risk and performance measurement systems and Basel II.

Nasrin R. Khalili, Ph.D., is an associate professor of Environmental Management at Illinois Institute of Technology, Stuart School of Business in Chicago. Dr. Khalili's research interest is in the areas of industrial pollution control, waste minimization, energy management, and environmental management system (EMS) design. She holds two patents and is the author of more than 35 referee articles and conference proceedings.

Dr. Khalili has extensive experience in working with industry on a wide range of pollution prevention, pollution control, waste minimization, and energy management projects. Since 1995, she has been collaborating in both research and education in the areas of environmental management with national and international universities such as RPI; NIU; UIC; School of Mining and Metallurgy in Krakow, Poland; Tecnológico de Monterrey, in Monterrey, Mexico; and the Foundation for Research and Technology in Environmental Management (FRTEM) in New Delhi, India.

Andrew Kumiega, Ph.D., has spent over 20 years automating processes, including CNC machining, chemical manufacturing, confectionary, pharmaceutical

manufacturing, and financial trading systems in industry as an industrial engineer. He has held various senior-level positions at financial institutions, including director of research at TD Waterhouse Securities Options; head of financial engineering at TFM Investments, LLC, and director of financial engineering at Market Liquidity Networks (all major options market makers); and vice president of quantitative research at Calamos Asset Management. Currently, he is employed at a proprietary trading firm. He is an adjunct professor at the Illinois Institute of Technology. He is a member of the American Society of Quality Control, a certified quality engineer, a certified quality auditor, and a certified software quality engineer. He is also a founding member of the market technology committee of the Certified Trading System Developer (CTSD) program at i4MT.

David Loshin is president of Knowledge Integrity, Inc. (www.knowledge-integrity.com), recognized worldwide as a thought leader in the areas of data quality, master data management, data governance, and business intelligence. David has contributed to many data management industry publications, including *Intelligent Enterprise, DM Review*, and *The Data Administration Newsletter* (www.tdan.com), and he currently is a channel expert at www.b-eye-network.com.

David's book *Business Intelligence: The Savvy Manager's Guide* (June 2003) has been hailed as a resource allowing readers to "gain an understanding of business intelligence, business management disciplines, data warehousing, and how all of the pieces work together." David's most recent book, *Master Data Management* (MK/OMG Press), has garnered endorsements from leaders across the data management industry, and his valuable MDM insights can be reviewed at www.mdmbook.com.

Michael Mainelli, Ph.D., FCCA FSI, originally undertook aerospace and computing research, followed by seven years as a partner in a large international accountancy practice, before a spell as corporate development director of Europe's largest R&D organization, the United Kingdom's Defence Evaluation and Research Agency, and becoming a director of Z/Yen (Michael_Mainelli@zyen.com). Z/Yen is the city of London's leading think tank, founded in 1994 in order to promote societal advance through better finance and technology. Z/Yen asks, solves, and acts globally on strategy, finance, systems, marketing and intelligence projects in a wide variety of fields (www.zyen.com), such as developing an award-winning risk/reward prediction engine, helping a global charity win a good governance award, or benchmarking transaction costs across global investment banks.

Z/Yen's humorous risk/reward management novel, *Clean Business Cuisine: Now and Z/Yen*, was published in 2000; it was a Sunday Times Book of the Week. *Accountancy Age* described it as "surprisingly funny considering it is written by a couple of accountants." Michael is Mercers' School Memorial Professor of Commerce at Gresham College.

Richard Marti, CISSP, CISA, QSA, is a principal at Computer Science Corporation (CSC) where he is building a Center of Excellence for Oracle GRC solutions. He is a subject matter expert for governance, risk, and compliance (GRC) solutions and has led multiple Sarbanes-Oxley (SOX), audit operations, IT governance, IT security, and compliance automation projects. He has been featured as a guest speaker on business and IT governance issues and has published papers on the Control

Objectives for Information and related Technology (COBIT)/Committee of Sponsoring Organizations of the Treadway Commission (COSO) framework, business continuity planning, and SOX compliance. He is contributor to two John Wiley & Sons texts by Anthony Tarantino: *Manager's Guide to Compliance* (March 2006) and *The Governance, Risk, and Compliance Handbook* (March 2008).

Bruce Rawlings is currently an independent consultant with trading and banking clients across the United States, with clients such as Mesirow, Advanced Strategies, and UBS Global Asset Management. He is an expert in Bayesian time series analysis with over 30 years in statistical modeling. Mr. Rawlings teaches graduate courses in econometrics, time series, quantitative investment strategies, interest rate modeling, and Bayesian econometrics at the Illinois Institute of Technology.

Claudio Schuster, CPA, CFE, and master in finance, has more than 25 years of experience in corporate finance and the financial markets in general. He also holds a management degree in energy from the University of Oxford. Claudio is a former VP at Citibank NA, Corporate Audit Division, and a chief financial officer at a major natural gas utility company in Argentina. During the Argentina debt crisis in 2001, Claudio was actively involved in the debt restructuring process. Presently, Claudio is the owner of The Financial People, a financial consultant firm, oriented to corporate finance and foreign exchange markets.

Brett Trusko, Ph.D., is a world-renowned Six Sigma Master Black Belt who has until recently led the process quality group for a major international consulting firm. His current position is as a quality researcher at the Medical College at Mayo Clinic. He is the author of hundreds of articles on quality and, as a futurist, has recently published a book, *Improving Healthcare Quality and Cost with Six Sigma*. He speaks and lectures globally on Six Sigma and his new approach, Dynamic Six Sigma. He has degrees in biology, accounting, and new product development, and a Ph.D. in information technology management.

Ben Van Vliet is a lecturer at the Illinois Institute of Technology's (IIT) Stuart School of Business, where he also serves as the associate director of the MS Financial Markets program. At IIT he teaches courses in quantitative finance, C++, and .NET programming, and automated trading system design and development. He is vice chairman of the Institute for Market Technology, where he chairs the advisory board for the Certified Trading System Developer (CTSD) program. He also serves as series editor of the Financial Markets Technology series for Elsevier/Academic Press. Mr. Van Vliet consults extensively in the financial markets industry, primarily on topics related to the mathematics, technology, and management of trading systems. He is the author of four books on trading/investment: *Quality Money Management* with Andrew Kumiega, *Modeling Financial Markets* with Robert Hendry, *Building Automated Trading Systems,* and *C++ with Financial Applications*. He has published several articles in the areas of finance and technology, and presented at several academic and professional conferences.

Chris Zephro is a director of finance for Seagate Technology, the largest manufacturer of hard disc drives. His extensive experience in Theory of Constraints includes implementation and training on the use of the TOC Thinking Process,

Constraint Exploitation using the Five Focusing Steps, and profit maximization leveraging throughput accounting. Chris has 15 years of experience in the field of supply chain management, operations, and finance; holds an MBA from the University of Tennessee; and has been practicing Theory of Constraints for over 12 years. He can be contacted at czephro@hotmail.com.

Risk Management in Finance

CHAPTER 1

Introduction

Anthony Tarantino, Ph.D., and Deborah Cernauskas, Ph.D.

Financial market turmoil is not a new phenomenon. From the tulip mania of the 1630s to the housing price bubble of the 2000s, the financial markets have been regularly subjected to periods of irrational behavior by investors and company management. The turmoil has not been confined to one country or geography and has been driven by various factors, including greed. Each period of turmoil creates many economic casualties, including lost jobs, corporate bankruptcies, and destroyed economic wealth.

Notwithstanding government regulations and oversight, financial turmoil and asset bubbles will continue to develop. The onus rightly lies with corporate executives and their boards of directors to act in the best interest of shareholders. Internal corporate oversight includes actively managing the risk-reward trade-off offered to shareholders. Corporate risk can take on many forms, including market, credit, and operational. The successful management and control of internal processes will increase the value of the firm by reducing operational losses and providing a competitive advantage. The focus of this book is on corporate management of internal processes generally classified as operational risk.

Operational risk is typically viewed as a risk arising from the execution of an organization's business functions. It has become a very broad concept, including risks from fraud, legal, physical, and environmental areas. Operational risk became a catch-all concept in financial institutions for any risk not credit or market related. Basel II is the capital accord developed for the banking industry by the Bank for International Settlements (BIS). Basel II defines operational risk as the risk of loss resulting from inadequate or failed internal processes, people, and systems, or from external events. Basel II has also created a classification for operational risk that is applicable to all industries. Basel II describes seven categories of operational risk:

1. *Internal Fraud*—misappropriation of assets, tax evasion, intentional mismarking of positions, bribery
2. *External Fraud*—theft of information, hacking damage, third-party theft, and forgery
3. *Employment Practices and Workplace Safety*—discrimination, workers' compensation, employee health and safety
4. *Clients, Products, and Business Practice*—market manipulation, antitrust, improper trade, product defects, fiduciary breaches, account churning

5. *Damage to Physical Assets*—natural disasters, terrorism, vandalism
6. *Business Disruption and Systems Failures*—utility disruptions, software failures, hardware failures
7. *Execution, Delivery, and Process Management*—data entry errors, accounting errors, failed mandatory reporting, negligent loss of client assets

In the past, high profit margins have characterized the financial services and banking industries. With the advent of commoditized Internet trading and banking services, the high profit margins are disappearing. The control of costs and risks are a high priority in a low-profit-margin environment.

Manufacturing firms have successfully dealt with quality control issues for many decades. Although the beginning of statistical process control is often accredited to Walter Shewhart who developed the control chart in 1924, the acceptance and use of process control did not occur until World War II, when wartime needs attached a high premium to product quality. After World War II, Japanese manufacturing went through a quality revolution. The quality focus shifted from product inspection to total process improvement. All organizational processes were subjected to quality improvements. The total quality initiative transformed Japanese manufacturing from a low-cost–low-quality producer to a low-cost–high-quality producer. By the end of the 1970s, Japan was the leading manufacturer of autos and electronics. The Toyota Production System, developed by Taiichi Ohno, became the basis of all subsequent just-in-time process improvements, which strive for the elimination of all waste. The United States responded to the Japanese total quality initiative with programs such as ISO 9000, Total Quality Management (TQM), Lean Manufacturing, and Six Sigma.

Over the past 40 years, statistical process control has been commonly implemented in the manufacturing, health care, and automotive industries through programs such as Six Sigma, and Lean Six Sigma. Six Sigma helps companies improve product quality and reduce waste by producing products and services better, cheaper, and faster.

The global financial crisis of 2007–2009 is only the latest example of economic turmoil caused by failures in financial risk management. The full extent of the economic, political, and human damage from the current crisis will not be known for some time, but it will dwarf the losses from Enron in the 1990s, the U.S. savings-and-loan crisis in the 1980s, and the Japanese banking crisis that occurred two decades ago.[1] The irony of the current crisis is that it occurred in an industry with the most sophisticated risk management systems and technologies and under very close government oversight. The current crisis is especially troubling in that risk management failed on multiple levels. At the most sophisticated level, quantitative and qualitative modeling gave few warnings of the huge risks inherent in leveraging capital at 30 to 1 and in assuming that real estate values would never decline. At the most simple level, common sense failed among investors, corporate executives and boards, rating agencies, and government regulators. Common sense should have warned that real estate values were growing at unsustainable rates, that middle-class folks were assuming far too much debt, and that making zero-down loans without verifying creditworthiness violated the most basic of banking practices.

Because of the depth and global reach of the current crisis, risk management is now an area of intense scrutiny far beyond corporate executives and government regulators. The demands for greater oversight and more robust risk management are

nearly universal. The pendulum has swung away from a laissez faire mentality with minimal market oversight to one in which regulators and stakeholders (investors, customers, suppliers, and community) will demand much tighter regulation. Unfortunately, greater regulation will fail unless coupled with much enhanced financial risk management. Regulators and corporate executives typically have a financial background but often lack financial risk management expertise. One could argue that the current crisis was the result of risk transparency failures, and not financial transparency failures. Increased risk transparency would help expose the dysfunctional nature of many operational risk management regimes.

ORGANIZATION OF THIS BOOK

The goal of this book is to provide an overview of some of the more exciting and effective techniques to improve financial risk management in operational areas. This is provided as a survey and not as an exhaustive treatment of every next-generation technique. We do cover the basics and include new and thought-provoking approaches that are applicable to all types and sizes of organizations, both public and private.

We begin with a survey of some of the foundations to financial risk management:

- Data Governance in Financial Risk Management
- Information Risk and Data Quality Control
- Total Quality Management
- Information Technology Risk
- Operational Risk Fundamentals
- Risk Management in Asia
- Risk Management in Latin America
- Risks in Migrating to the International Financial Reporting Standards (IFRS)
- Quantitative Operational Risk Methods

We follow with next-generation best practices to improve financial risk management:

- Statistical Process Control Integrated with Engineering Process Control
- Business Process Management Integrated with Lean Six Sigma
- Bayesian Networks for Root Cause Analysis
- Information Analytics
- Embedded Predictive Analytics
- Reducing Risk in Litigation and Legal Discovery
- The Circle of Trust
- Reducing Risk with Environmental Best Practices
- Next-Generation Techniques in Segregation of Duties
- Transaction Based Cross-Enterprise Risk Management
- Throughput Accounting
- Environmental Consistency Confidence
- Quality in the Front Office—Reducing Process Variation in Trading Firms
- Root Cause of the Global Financial Crisis and Corporate Governance Reforms to Prevent the Next Failure in Risk Management

WHY READ THIS BOOK?

The goal of this book is to aid financial professionals in implementing quality assurance systems for financial processes that will in turn enable data-driven decision making. The catastrophic failures of risk management behind the global financial crisis demonstrate the criticality of improving the quality and risk management processes in financial services.

The stakes are extremely high—the laggards are doomed to continue to suffer through enterprise-threatening risk failures. The leaders will never be free of risk failures, but will substantially increase their ability to successfully balance risk and reward opportunities.

NOTE

1. Carrick Mollenkamp and Mark Whitehouse, "Banks Fear a Deepening of Turmoil," *Wall Street Journal*, March 17, 2008, pp. 1, 12.

CHAPTER 2

Data Governance in Financial Risk Management

Anthony Tarantino, Ph.D.

INTRODUCTION

Let's start with a definition of governance and data governance. *Governance* is the act of governing or exercising authority over those who are governed by persons and organizations who are part of a body that has the responsibility for administering something. *Data governance* is simply the governance of the people, process, and technology applied to data used by an organization to ensure its definition, validity, consistency, quality, timeliness, and availability to the appropriate owners and users of the data. For our purposes, "data is any information captured within a computerized system, which can be represented in graphical, text or speech form."[1]

Complicating data governance is the issue of paper documents. In today's organizations, it is rare for paper documents not to originate in some sort of electronic or digital format. This is becoming a major issue in litigation and regulatory audits. Litigants, regulators, and auditors are less and less willing to accept paper documents without electronic metadata references as to ownership, access and change controls, time stamps, and so on. The reason is simple: it is very easy to fake a paper document. So, by extension, data governance is not just over digital data, but all data—paper and electronic.

Data governance is not the same as data management. Data management is a subcomponent of data governance and includes the management of data and metadata access points. Documents and records management, often referred to as enterprise content management (ECM), can be seen as a subset of data governance as well and includes the technologies used to capture, manage, store, preserve, and deliver content and documents related to organizational processes.[2] ECM is typically a process to control unstructured data, while data governance controls all types of data—structured, semistructured, unstructured, metadata, registries, ontologies, and taxonomies.[3]

Unstructured data creates headaches for most all organizations in achieving data governance. Even its definition is debatable. Unstructured data is typically said to be data that is not readily readable by computers, such as e-mails, instant messages, word processer documents, audio, and video. It typically represents the great majority of all data in any organization, and the trend is accelerating with the

growth of instant messages and e-mails. Data with some type of structure may also be classified as unstructured if its structure does not support the needed processing task. For instance, while an HTML (hypertext markup language) web page is tagged, the tag is to support its format and not its meaning.[4]

And why is data governance so critical in financial risk management? Simply put, data and its management are key in all organizations. Without very robust controls over data, an organization is exposed to high levels of financial risk. Today's financial institutions, including banks, excel when they move the right data at the right time to the right users of data. Nonfinancial institutions also rely on robust data governance to prosper. Health care enterprises worry about patient data and maintaining its privacy. Pharmaceutical enterprises worry about documenting their compliance with complex regulations. Manufacturing and distribution companies worry about inventory and bills-of-material accuracy, retailers worry about capturing point of sales in real time. All firms worry about consolidating financial information to their general ledgers and to support period-end closes and audits.

The importance of data governance is not a new concept. Dating back to 1500 B.C., the Phoenicians built an empire based on trade and commerce. This required a system of mass communication for accurate record keeping and streamlined communication. It began as a cuneiform system of characters developed in Mesopotamia and evolved into the world's first alphabet, needed for more accurate and mobile record keeping. Registry filing systems date back to ancient Rome, survive today in many parts of the world, and represent a best practice in early record-keeping systems. Officials maintained *commentarii,* or private notes, which they consolidated daily into court journals, or *commentarii diarni.* These journal entries were maintained for all inbound and outbound types of documents, including court rulings, litigations, and contract transactions.[5] The Phoenicians, Romans, and other ancients well understood the criticality of data governance and the major risks when data governance failed. The proof can be found in the amazingly detailed records that have survived for the most minor of commercial, government, and military activities and transactions. The main difference is the huge amounts and many types of data that must be maintained in real time today.

DATA GOVERNANCE CENTER OF EXCELLENCE

An essential first step in achieving data governance (DG) is to create a center of excellence (CoE) around it. Some have called for a data governance council as a central focal point of DG activity, but a DG CoE takes this beyond a bureaucratic organization that merely coordinates activities to a group that owns and communicates the organization's vision of DG. Without a CoE, an organization may have a different vision for each of its lines of business, regions, and/or information technology (IT) environments. A DG CoE should be involved with the following activities:

- It fully understands the organization's current state of DG. This includes periodic surveys of all lines of business, locations, and IT environments.
- It develops a desired DG end state based on the desires and business requirements of all the organization's DG stakeholders. The desired end state is approved by the organization's executive management, external auditors, and applicable

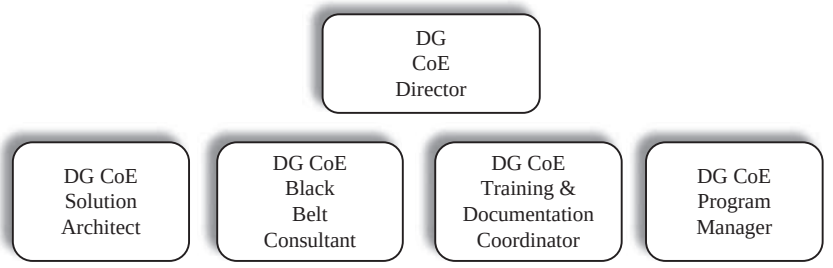


EXHIBIT 2.1 Data Governance Center of Excellence Organization Chart

regulatory agencies. Once approved, the desired end state is communicated to the entire organization and its stakeholders.
- It coordinates periodic DG assessments, which include a current state, desired end state, gap analysis, and cost-benefit analysis. This is more fully described in the next section.
- It reviews, coordinates, and approves all enterprise-wide DG guidelines, policies, procedures, audit procedures, risk-control matrices, and workflows. This is not to say that they usurp local controls, only that they provide oversight that captures the organization's DG vision.
- It strives to eliminate disparate DG practices and move the organization to enterprise-wide practices based on industry-accepted best practice frameworks.

The DG CoE should include representatives of each line of business, IT, legal, and internal audit. It need not be a large organization and can include only a small dedicated staff that could look something like Exhibit 2.1 in its initial phases.

- *DG CoE Director* is responsible for championing the organization's DG vision and coordinating all significant DG initiatives across the organization. This includes the communication of critical activities and issues to the executive management, auditors, and legal counsel; facilitating required DG structures; and coordinating enterprise-wide DG architecture development plans and support requirements.
- *DG CoE Solution Architect* ensures that the agreed-upon technical architectures and standards are communicated and adhered to across the organization. This includes providing program and project oversight and coordination, and developing and communicating new processes and best practices.
- *DG CoE Black Belt* applies proven Six Sigma process improvement and problem-solving techniques to attack the most significant DG problems the organization faces. Black belts strive to respond to the voice of the customer—both internal and external customers—and to reduce variability in a given process. The result is higher-quality processes and lower financial risk. They act as an internal consultant to support all the lines of business, with their priorities set by the DG CoE Director. Many black belts are also trained in Lean processes pioneered by Toyota back in the 1960s and 1970s. Lean Six Sigma combines the strengths of both philosophies.

- *DG CoE Training and Documentation Coordinator* promotes education and training in DG procedures and guidelines. This includes maintaining and communicating the relevant training materials; tracking acceptance and acceptance issues to DG procedures and guidelines; and assuring the quality, consistency, and availability of the training process.
- *DG CoE Program Director* oversees all relevant DG projects and programs (multiple projects with interrelated objectives and dependent tasks). This includes tracking and communicating their status, resource staffing, critical issues, actual costs to budgeted costs, and dependencies.

DATA GOVERNANCE ASSESSMENT

For an organization to understand its DG current state, and gaps to achieve its desired end state, it is helpful to conduct an assessment. This is a traditional process in problem solving widely used by consultants and process improvement teams.

It begins by capturing the current state of DG across the enterprise. This is typically no minor task in decentralized organizations with heterogeneous IT environments and multiple silos of data in which many practices are not documented or are poorly understood outside of the business units and geographic locations. It is important to capture both the strengths and weaknesses, as islands of strengths can be used as role models for the rest of the organization.

Next, it is necessary to survey the business owners as to how they would define DG success. Of course, it is unlikely that there will be a great deal of consistency in their definition of success and the desired end state. It makes sense to first charter a DG CoE to take ownership of defining the desired end state. The alternative will be to present a variety of disparate and confusing ideas to an organization's executive management. The desired end state should not be made in isolation but leverage best practice frameworks such as Control Objectives for Information and related Technology (COBIT), Information Technology Infrastructure Library (ITIL), National Institute of Standards and Technology (NIST) 800, and related International Organization for Standardization (ISO) standards. There is no need to start with a blank sheet.

Once the desired end state is agreed upon, the next step is to perform a gap analysis. The gap analysis should incorporate the risks of doing nothing and the risks, costs, and benefits of closing the gaps.

The final phase is to prepare a proposed action plan to achieve the end state including a prioritization of each objective. Achieving best practices and next-generation techniques in DG is a daunting task. Some goals will take years to achieve, while others are fairly short term. Overwhelming an organization with unattainable or excessive stretch goals will backfire and create more problems than will doing nothing.

DATA GOVERNANCE MATURITY MODEL

The assessment process can be enhanced by rating the organization against a data governance maturity model (see Exhibit 2.2). In this model, the least mature

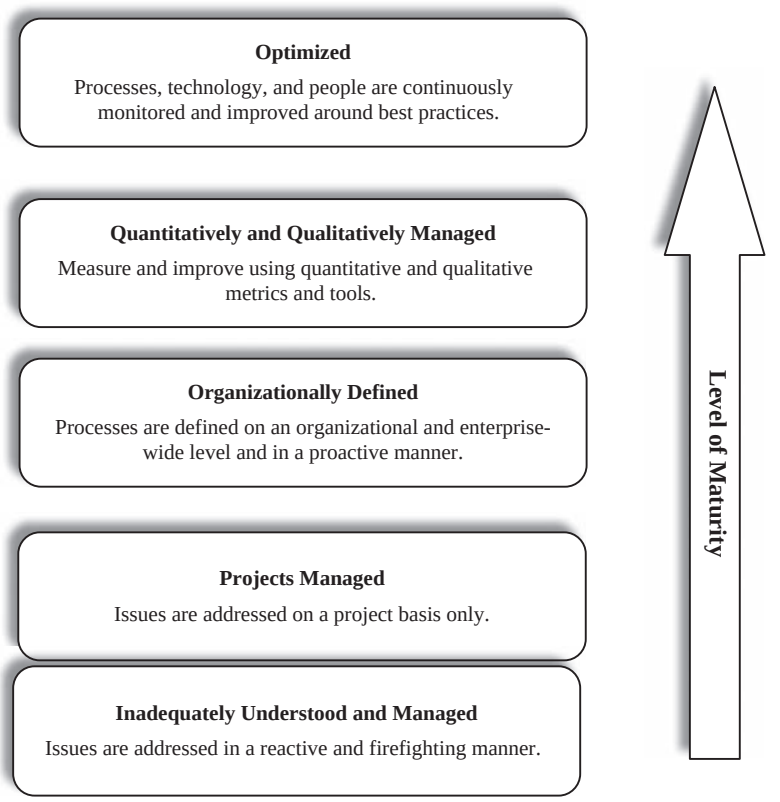


EXHIBIT 2.2 Data Governance Maturity Model

organizations are in a reactive and firefighting mode. As organizations improve, they begin to move from a project to an enterprise-wide approach. Ultimately, they use qualitative and quantitative metrics to continuously monitor and improve their people, processes, and technologies.

The unfortunate reality is that many organizations are at the lowest levels of the maturity model. These are some of the characteristics to look for in an organization that is challenged by its DG:

- *Data quality.* Data governance ownership and accountability are not clearly defined, understood, or adhered to. Enterprise-wide policies, procedures, guidelines, and standards are lacking. Data governance is viewed by business owners and stakeholders as an IT issue. IT addresses DG in application and business silos.
- *Data architecture.* An enterprise-wide data architecture is not in place and each application and database owner has their own definition of data and applicable standards. There is typically little sharing of data or efforts to find a common framework.

- *General IT environment*. The IT infrastructure is overly complex, applications are silo driven, data accuracy is typically inconsistent in and across the lines of business, and IT initiatives are sometimes redundant and poorly coordinated.
- *Metadata*. There is a lack of consistency and standardization in the collection and storage of metadata. There is no enterprise-wide program to associate all digital data upon creation to its applicable metadata.
- *Policies and procedures*. There is no viable system of policies and procedures in force to control the data governance process. As a consequence, activities are reactive and ad hoc.
- *Security and privacy*. There is a lack of adherence to accepted best practice standards in security and privacy protection.
- *Information life-cycle management*. While there are some policies in place around data retention and destruction, enforcement is inconsistent and not well understood.
- *Tone at the top*. The organization understands the basics of the regulatory, risk, and legal discovery drivers behind data governance, but lacks the executive sponsorship (or tone at the top) to instill the critical importance of data organization to the well-being and survival of the organization.

In April 2006, International Business Machines (IBM) sponsored a survey of the current state and best practices in data governance among 50 Global 500 organizations.[6] A summary of findings demonstrates the major challenges most face in improving data governance:

- Only about one quarter of firms enjoy central data ownership.
- Only one half have key performance indicators and metrics that define DG success.
- Only one third have defined and communicated to the organization their DG program (objectives, goals, milestones, executive ownership, etc.).

BEST PRACTICES IN DATA GOVERNANCE

These are some techniques that will help improve DG regardless of the industry, IT environment, and complexity of data:

- *Determine the value and risk of the data*. Because organizations need to address DG from a wide variety of sources, it is helpful to prioritize data as to its value to the organization. Once its value is determined, the next step is to calculate the risks associated with it. Once its value and risk are determined, it is now possible to determine what to budget in terms of finances and resources to manage it.
- *Digitize all content upon origination*. Given the masses of disparate data that all organizations must address, it is critical to digitize all data upon origination. This includes two critical steps:
  1. *Classify and index data to its metadata references*. This tags all data upon origination as to ownership, date of creation, revision, or access, and its nature. Without this, data is not easily searched or accessed, making for a painfully expensive and tedious audit and legal discovery process.

2. *Destroy all paper originals once they have been digitized.* Unless prohibited by regulatory requirements, paper originals create an undue burden on an organization. The acceptance of digital signatures is now commonplace, saving the expense and physical space to maintain paper documents and records. Paper documents are particularly a burden in the legal discovery process in which litigants demand to the see the electronic metadata references to all paper documents. In short, a piece of paper has little value unless it can be tied to ownership, access controls, time stamp, and chain of possession.

- *Reduce the number of content repositories.* Data governance is simplified with the reduction of the number of data repositories and the standardization of the data in those that remain. Reducing content repositories also has the benefit of compelling a standardization of the formats and naming conventions as repositories are eliminated. In an age of ongoing mergers, acquisitions, and consolidations, it is typical to find a wide variety of DG standards in place. For unstructured data, this can translate to the same customer, supplier, or item listed under a wide variety of names and classification codes—none of which are easily discovered by the organization.
- *Federate content across repositories.* Unfortunately, we live in a very heterogeneous IT environment for most organizations where it is not cost effective or even possible to eliminate multiple data repositories. Federation of content provides the means to access multiple data repositories and in effect create a virtual data repository. More complex federation permits cross-referencing and accessing all documents and records that are related, such as all records related to a given customer or supplier. For example, with complex federated content, a bank would be able to easily access a customer's savings, checking, credit card, retirement account, car loan, and home loan—even if each exists under separate lines of business in a separate databases.
- *Expand the use of WORM technology.* Write once and read many times (WORM) technology is widely available in optical, disc, and tape formats. WORM technology helps to assure that there is no unauthorized or undocumented update to protected documents and records.
- *Expand the use of business process management (BPM) and workflows.* Electronic workflows have been readily available for several years and can go a long way in improving the DG process. When combined with Lean Six Sigma process improvement (the topic of Chapter 12), they offer a next-generation technique in streamlining and automating processes and the data associated with those processes. Typically, BPM includes automating tasks, approvals, and forms, which provides for a transparent and end-to-end audit trail—highly desirable to auditors, regulators, and risk managers. The nature of BPM facilitates standardizing processes and, when combined with Lean and Six Sigma, will help to standardize around optimized best practices. DG is bound to improve with the used of automated workflows, approvals, and electronic forms that have been standardized on an enterprise level.
- *Expand the use of data quality tools.* Data quality tools compare data against a data quality standard. Outputs can include the identification of duplicated master level (supplier, customer, item, commodity code, etc.). Some commodity coding tools will attempt to assign the proper code based on item descriptions. The problem arises in that any given item can be described in many ways.

Commodity codes such as the United Nations Standard Products and Services Classification (UN/SPSC) have helped to standardize the commodity coding process with an open, global, and hierarchical standard.[7] For example, what commodity code should be used for an office trash bin? Is it an office supply, janitorial supply, or storage container, to name a few? Without an accepted standard, it is not unusual to find the same people applying a variety of different commodity codes to the same types of items. It is also not unusual to see suppliers and customers duplicated, even when only one individual owns the process. I recall a one-person procurement and accounts payable department listing their supplier, Owens Corning, under at least four different names: Owens Corning International, Owens-Corning Inc., Owens-Corning International at a corporate P.O. box, and Owens Corning at a local address—each with a different supplier number. Data quality tools will help to identify such obvious duplications. Eliminating them is typically not an easy process and requires cross-referencing and merging of histories.

CONCLUSION: NEXT-GENERATION TECHNIQUES TO REDUCE DATA GOVERNANCE RISK

It is well understood that DG is vital in most organizations, especially those that are heavily regulated, those subject to ever more demanding regulatory audits, and those in highly litigious environments in which legal discovery is now a major cost of doing business. Many of the best practices we describe are in common practice, but the next-generation techniques that will provide a strategic competitive advantage are not in the planning stages in most organizations.

Next-generation DG will require the formation and full executive support for a DG CoE, not just a council of department heads. The DG CoE must provide a strategic vision for DG that supports the organization's objectives. It must also act as the clearinghouse, facilitator, coordinator, and program overseer for all related DG initiatives.

Staffing a DG CoE with Lean Six Sigma black belts who champion the voice of the customer and reduce variability in a process will change the focus of DG from an IT-driven/owned process to one that strives to meet all customer expectations. These customers include other internal departments, auditors, regulators, and suppliers, as well as external customers.

A DG self-assessment is usually mentioned as a first step in any improvement process. But without first forming a DG CoE, a DG self-assessment is likely to fail in understanding the organization's DG environment and developing the most viable desired future/end state. Without a DG CoE, the DG projects funded from the initial assessment will be less effective, with greater redundancies and gaps. Finally, a DG CoE will help prioritize the many competing DG initiatives and help select the most appropriate best practices.

The management and governance of data presents a major risk and opportunity to all organizations. Well-managed and governed data is a major asset, but poorly managed and governed data represents major liabilities. The massive growth in unstructured data will continue to compound the problems organizations face. The recommendations made here are no panacea, but can help provide a strategic and

competitive advantage over those who treat DG in an ad hoc manner and assign its ownership to their IT departments.

NOTES

1. Anne Marie Smith, "Data Governance Best Practices—The Beginning," EIMInstitute.org, www.eiminstitute.org/library/eimi-archives/volume-1-issue-1-march-2007-edition/data-governance-best-practices-2013-the-beginning.
2. See The Association for Information and Image Management (AIIM) web site: www.aiim.org/.
3. See note 1.
4. Ibid.
5. David Stephens, "Registry: The World's Most Predominant Recordkeeping System," *ARMA Records Management Quarterly,* January 1995.
6. CDI Institute, "Corporate Data Governance Best Practices: 2006–07 Scorecards for Data Governance in the Global 5000," April 2006.
7. See www.unspsc.org/.

CHAPTER 3

Information Risk and Data Quality Management

David Loshin

INTRODUCTION

It would not be a stretch of the imagination to claim that most organizations today are heavily dependent on the use of information to both *run* and *improve* the ways that they achieve their business objectives. That being said, the reliance on dependable information introduces risks to the ability of a business to achieve its business goals, and this means that no enterprise risk management program is complete without instituting processes for assessing, measuring, reporting, reacting to, and controlling the risks associated with poor data quality.

However, the consideration of information as a fluid asset, created and used across many different operational and analytic applications, makes it difficult to envision ways to assess the risks related to data failures as well as ways to monitor conformance to business user expectations. This requires some exploration into types of risks relating to the use of information, ways to specify data quality expectations, and developing a data quality scorecard as a management tool for instituting data governance and data quality control.

In this chapter we look at the types of risks that are attributable to poor data quality as well as an approach to correlating business impacts to data flaws. Data governance (DG) processes can contribute to the description of data quality expectations and the definition of relevant metrics and acceptability thresholds for monitoring conformance to those expectations. Combining the raw metrics scores with measured staff performance in observing data service-level agreements contributes to the creation of a data quality scorecard for managing risks.

ORGANIZATIONAL RISK, BUSINESS IMPACTS, AND DATA QUALITY

If successful business operations rely on high-quality data, then the opposite is likely to be true as well: flawed data will delay or obstruct the successful completion of business processes. Determining the specific impacts that are related to the different data issues that emerge is a challenging process, but assessing impact is simplified

through the characterization of impacts within a business impact taxonomy. Categories in this taxonomy relate to aspects of the business's financial, confidence, and compliance activities, yet all business impact categories deal with enterprise risk. There are two aspects of looking at information and risk; the first looks at how flawed information impacts organizational risk, while the other looks at the types of data failures that create the exposure.

Business Impacts of Poor Data Quality

Many data quality issues may occur within different business processes, and a data quality analysis process should incorporate a business impact assessment to identify and prioritize risks. To simplify the analysis, the business impacts associated with data errors can be categorized within a classification scheme intended to support the data quality analysis process and help in distinguishing between data issues that lead to material business impact and those that do not. This classification scheme defines six primary categories for assessing either the negative impacts incurred as a result of a flaw, or the potential opportunities for improvement resulting from improved data quality:

1. Financial impacts, such as increased operating costs, decreased revenues, missed opportunities, reduction or delays in cash flow, or increased penalties, fines, or other charges.
2. Confidence-based impacts, such as decreased organizational trust, low confidence in forecasting, inconsistent operational and management reporting, and delayed or improper decisions.
3. Satisfaction impacts such as customer, employee, or supplier satisfaction, as well as general market satisfaction.
4. Productivity impacts such as increased workloads, decreased throughput, increased processing time, or decreased end-product quality.
5. Risk impacts associated with credit assessment, investment risks, competitive risk, capital investment and/or development, fraud, and leakage.
6. Compliance is jeopardized, whether that compliance is with government regulations, industry expectations, or self-imposed policies (such as privacy policies).

Despite the natural tendency to focus on financial impacts, in many environments the risk and compliance impacts are largely compromised by data quality issues. Some examples to which financial institutions are particularly sensitive include:

- Anti-money laundering aspects of the Bank Secrecy Act and the USA PATRIOT Act have mandated private organizations to take steps in identifying and preventing money laundering activities that could be used in financing terrorist activities.
- Sarbanes-Oxley, in which section 302 mandates that the principal executive officer or officers and the principal financial officer or officers certify the accuracy and correctness of financial reports.
- Basel II Accords provide guidelines for defining the regulations as well as guiding the quantification of operational and credit risk as a way to determine the

amount of capital financial institutions are required to maintain as a guard against those risks.
- The Graham-Leach-Bliley Act of 1999 mandates financial institutions with the obligation to "respect the privacy of its customers and to protect the security and confidentiality of those customers' nonpublic personal information."
- Credit risk assessment, which requires accurate documentation to evaluate an individual's or organization's abilities to repay loans.
- System development risks associated with capital investment in deploying new application systems emerge when moving those systems into production is delayed due to lack of trust in the application's underlying data assets.

While the sources of these areas of risk differ, an interesting similarity emerges: not only do these mandate the use or presentation of high-quality information, they also require means of demonstrating the adequacy of internal controls overseeing that quality to external parties such as auditors. This means that not only must financial institutions manage the quality of organizational information, they must also have governance processes in place that are transparent and auditable.

Information Flaws

The root causes for the business impacts are related to flaws in the critical data elements upon which the successful completion of the business processes depend. There are many types of erred data, although these common issues lead to increased risk:

- Data entry errors
- Missing data
- Duplicate records
- Inconsistent data
- Nonstandard formats
- Complex data transformations
- Failed identity management processes
- Undocumented, incorrect, or misleading metadata

All of these types of errors can lead to inconsistent reporting, inaccurate aggregation, invalid data mappings, incorrect product pricing, and failures in trade settlement, among other process failures.

EXAMPLES

The general approach to correlating business impacts to data quality issues is not new, and in fact there are some interesting examples that demonstrate different types of risks that are attributable to flaws (both inadvertent and deliberate) in data.

Employee Fraud and Abuse

In 1997, the Department of Defense Guidelines on Data Quality categorized costs into four areas: prevention, appraisal, internal failure, and external failure. In turn, the impacts were evaluated to assess costs to correct data problems as opposed to

costs incurred by ignoring them. Further assessment looked at direct costs (such as costs for appraisal, correction, or support) versus indirect costs (such as customer satisfaction). That report documents examples of how poor data quality impacts specific business processes: "... the inability to match payroll records to the official employment record can cost millions in payroll overpayments to deserters, prisoners, and 'ghost' soldiers. In addition, the inability to correlate purchase orders to invoices is a major problem in unmatched disbursements."[1]

The 2006 Association of Certified Fraud Examiners Report to the Nation[2] details a number of methods that unethical employees can use to modify existing data to commit fraudulent payments. Invalid data is demonstrated to have significant business impacts, and the report details median costs associated with these different types of improper disbursements.

Underbilling and Revenue Assurance

NTL, a cable operator in the United Kingdom, anticipated business benefits in improving the efficiency and value of an operator's network through data quality improvement. Invalid data translated into discrepancies between services provided and services invoiced, resulting in a waste of unknown excess capacity. Their data quality improvement program was, to some extent, self-funded through the analysis of "revenue assurance to detect under billing. For example, ... results indicated leakage of just over 3 percent of total revenue."[3]

Credit Risk

In 2002, a PriceWaterhouseCoopers study on credit risk data indicated that a significant percentage of the top banks were deficient in credit risk data management, especially in the areas of counterparty data repositories, counterparty hierarchy data, common counterparty identifiers, and consistent data standards.[4]

Insurance Exposure

A 2008 Ernst & Young survey on catastrophe exposure data quality highlighted that "shortcomings in exposure data quality are common," and that "not many insurers are doing enough to correct these shortcomings," which included missing or inaccurate values associated with insured values, locations, building class, occupancy class, as well as additional characteristics.[5]

Development Risk

Experience with our clients has indicated a common pattern in which significant investment in capital acquisitions and accompanying software development has been made in the creation of new business application systems, yet the deployment of those systems is delayed (or perhaps even canceled) due to organizational mistrust of the application data. Such delayed application development puts investments at risk.

Compliance Risk

Pharmaceutical companies are bound to abide by the federal Anti-Kickback Statute, which restricts companies from offering or paying remuneration in return for

arranging for the furnishing of items or services for which payment may be made under Medicare or a state health care program. Pharmaceutical companies fund research using their developed products as well as market those same products to potentially the same pool of practitioners and providers, so there is a need for stringent control and segregation of the data associated with both research grants and marketing.

Our experience with some of our clients has shown that an assessment of party information contained within master data sets indicated some providers within the same practice working under research grants while others within the same practice subjected to marketing. Despite the fact that no individual appeared within both sets of data, the fact that individuals rolled up within the same organizational hierarchy exposed the organization to potential violation of the Anti-Kickback Statute.

DATA QUALITY EXPECTATIONS

These examples are not unique, but instead demonstrate patterns that commonly emerge across all types of organizations. Knowledge of the business impacts related to data quality issues is the catalyst to instituting data governance practices that can oversee the control and assurance of data validity. The first step toward managing the risks associated with the introduction of flawed data into the environment is articulating the business user expectations for data quality and asserting specifications that can be used to monitor organizational conformance to those expectations. These expectations are defined in the context of "data quality dimensions," high-level categorizations of assertions that lend themselves to quantification, measurement, and reporting.

The intention is to provide an ability to characterize business user expectations in terms of acceptability thresholds applied to quantifiers for data quality that are correlated to the different types of business impacts, particularly the different types of risk. And although the academic literature in data quality enumerates many different dimensions of data quality, an initial development of a data quality scorecard can rely on a subset of those dimensions, namely, accuracy, completeness, consistency, reasonableness, currency, and identifiability.

Accuracy

The dimension of accuracy measures the degree with which data instances compare to the "real-life" entities they are intended to model. Often, accuracy is measured in terms of agreement with an identified reference source of correct information such as a "system of record," a similar corroborative set of data values from another table, comparisons with dynamically computed values, or the results of manually checking value accuracy.

Completeness

The completeness dimension specifies the expectations regarding the population of data attributes. Completeness expectations can be measured using rules relating to

varying levels of constraint—mandatory attributes that require a value, data elements with conditionally optional values, and inapplicable attribute values.

Consistency

Consistency refers to measuring reasonable comparison of values in one data set to those in another data. Consistency is relatively broad, and can encompass an expectation that two data values drawn from separate data sets must not conflict with each other, or define more complex comparators with a set of predefined constraints. More formal consistency constraints can be encapsulated as a set of rules that specify relationships between values of attributes, either across a record or message, or along all values of a single attribute.

However, be careful not to confuse consistency with accuracy or correctness. Consistency may be defined between one set of attribute values and another attribute set within the same record (record-level consistency), between one set of attribute values and another attribute set in different records (cross-record consistency), or between one set of attribute values and the same attribute set within the same record at different points in time (temporal consistency).

Reasonableness

This dimension is used to measure conformance to consistency expectations relevant within specific operational contexts. For example, one might expect that the total sales value of all the transactions each day is not expected to exceed 105 percent of the running average total sales for the previous 30 days.

Currency

This dimension measures the degree to which information is current with the world that it models. Currency measures whether data is considered to be "fresh," and its correctness in the face of possible time-related changes. Data currency may be measured as a function of the expected frequency rate at which different data elements are expected to be refreshed, as well as verifying that the data is up to date. Currency rules may be defined to assert the "lifetime" of a data value before it needs to be checked and possibly refreshed.

Uniqueness

This dimension measures the number of inadvertent duplicate records that exist within a data set or across data sets. Asserting uniqueness of the entities within a data set implies that no entity exists more than once within the data set and that there is a key that can be used to uniquely access each entity (and only that specific entity) within the data set.

Other Dimensions of Data Quality

This list is by no means complete—there are many other aspects of expressing the expectations for data quality, such as semantic consistency (dealing with the consistency of meanings of data elements), structural format conformance, timeliness, and

valid ranges, valid within defined data domains, among many others. The principal concept is that the selected dimensions characterize aspects of the business user expectations and that they can be quantified using a reasonable measurement process.

MAPPING BUSINESS POLICIES TO DATA RULES

Having identified the dimensions of data quality that are relevant to the business processes, we can map the information policies and their corresponding business rules to those dimensions. For example, consider a business policy that specifies that personal data collected over the web may be shared only if the user has not opted out of that sharing process. This business policy defines information policies: the data model must have a data attribute specifying whether a user has opted out of information sharing, and that attribute must be checked before any records may be shared. This also provides us with a measurable metric: the count of shared records for those users who have opted out of sharing.

The same successive refinement can be applied to almost every business policy and its corresponding information policies. As we distill out the information requirements, we also capture assertions about the business user expectations for the result of the operational processes. Many of these assertions can be expressed as rules for determining whether a record does or does not conform to the expectations. The assertion is a quantifiable measurement when it results in a count of nonconforming records, and therefore monitoring data against that assertion provides the necessary data control.

Once we have reviewed methods for inspecting and measuring against those dimensions in a quantifiable manner, the next step is to interview the business users to determine the acceptability thresholds. Scoring below the acceptability threshold indicates that the data does not meet business expectations, and highlights the boundary at which noncompliance with expectations may lead to material impact to the downstream business functions. Integrating these thresholds with the methods for measurement completes the construction of the data quality control. Missing the desired threshold will trigger a data quality event, notifying the data steward and possibly even recommending specific actions for mitigating the discovered issue.

DATA QUALITY INSPECTION, CONTROL, AND OVERSIGHT: OPERATIONAL DATA GOVERNANCE

In this section we highlight the relationship between data issues and their downstream impacts, and note that being able to control the quality of data throughout the information processing flow will enable immediate assessment, initiation of remediation, and an audit trail demonstrating the levels of data quality as well as the governance processes intended to ensure data quality.

Operational data governance is the manifestation of the processes and protocols necessary to ensure that an acceptable level of confidence in the data effectively satisfies the organization's business needs. A data governance program defines the roles, responsibilities, and accountabilities associated with managing data quality. Rewarding those individuals who are successful at their roles and responsibilities can ensure the success of the data governance program. To measure this, a "data quality

scorecard" provides an effective management tool for monitoring organizational performance with respect to data quality control.

Operational data governance combines the ability to identify data errors as early as possible with the process of initiating the activities necessary to address those errors to avoid or minimize any downstream impacts. This essentially includes notifying the right individuals to address the issue and determining if the issue can be resolved appropriately within an agreed-to time frame. Data inspection processes are instituted to measure and monitor compliance with data quality rules, while service-level agreements (SLAs) specify the reasonable expectations for response and remediation.

Note that data quality inspection differs from data validation. While the data validation process reviews and measures conformance of data with a set of defined business rules, inspection is an ongoing process to:

- Reduce the number of errors to a reasonable and manageable level.
- Enable the identification of data flaws along with a protocol for interactively making adjustments to enable the completion of the processing stream.
- Institute a mitigation or remediation of the root cause within an agreed-to time frame.

The value of data quality inspection as part of operational data governance is in establishing trust on behalf of downstream users that any issue likely to cause a significant business impact is caught early enough to avoid any significant impact on operations. Without this inspection process, poor-quality data pervades every system, complicating practically any operational or analytical process.

MANAGING INFORMATION RISK VIA A DATA QUALITY SCORECARD

While there are practices in place for measuring and monitoring certain aspects of organizational data quality, there is an opportunity to evaluate the relationship between the business impacts of noncompliant data as indicated by the business clients and the defined thresholds for data quality acceptability. The degree of acceptability becomes the standard against which the data is measured, with operational data governance instituted within the context of measuring performance in relation to the data governance procedures. This measurement essentially covers conformance to the defined standards, as well as monitoring staff agility in taking specific actions when the data sets do not conform. Given the set of data quality rules, methods for measuring conformance, the acceptability thresholds defined by the business clients, and the SLAs, we can monitor data governance by observing not only compliance of the data to the business rules, but of the data stewards to observing the processes associated with data risks and failures.

The dimensions of data quality provide a framework for defining metrics that are relevant within the business context while providing a view into controllable aspects of data quality management. The degree of reportability and controllability may differ depending on one's role within the organization, and correspondingly, so will the level of detail reported in a data quality scorecard. Data stewards may

focus on continuous monitoring in order to resolve issues according to defined SLAs, while senior managers may be interested in observing the degree to which poor data quality introduces enterprise risk.

Essentially, the need to present higher-level data quality scores introduces a distinction between two types of metrics. The simple metrics based on measuring against defined dimensions of data quality can be referred to as "base-level" metrics, and they quantify specific observance of acceptable levels of defined data quality rules. A higher-level concept would be the "complex" metric representing a rolled-up score computed as a function (such as a sum) of applying specific weights to a collection of existing metrics, both base-level and complex. The rolled-up metric provides a qualitative overview of how data quality impacts the organization in different ways, since the scorecard can be populated with metrics rolled up across different dimensions depending on the audience. Complex data quality metrics can be accumulated for reporting in a scorecard in one of three different views: by **issue,** by **business process,** or by **business impact.**

Data Quality Issues View

Evaluating the impacts of a specific data quality issue across multiple business processes demonstrates the diffusion of pain across the enterprise caused by specific data flaws. This scorecard scheme, which is suited to data analysts attempting to prioritize tasks for diagnosis and remediation, provides a rolled-up view of the impacts attributed to each data issue. Drilling down through this view sheds light on the root causes of impacts of poor data quality, as well as identifying "rogue processes" that require greater focus for instituting monitoring and control processes.

Business Process View

Operational managers overseeing business processes may be interested in a scorecard view by business process. In this view, the operational manager can examine the risks and failures preventing the business process's achievement of the expected results. For each business process, this scorecard scheme consists of complex metrics representing the impacts associated with each issue. The drill-down in this view can be used for isolating the source of the introduction of data issues at specific stages of the business process as well as informing the data stewards in diagnosis and remediation.

Business Impact View

Business impacts may have been incurred as a result of a number of different data quality issues originating in a number of different business processes. This reporting scheme displays the aggregation of business impacts rolled up from the different issues across different process flows. For example, one scorecard could report rolled-up metrics documenting the accumulated impacts associated with credit risk, compliance with privacy protection, and decreased sales. Drilling down through the metrics will point to the business processes from which the issues originate; deeper review will point to the specific issues within each of the business processes. This view is suited to a more senior manager seeking a high-level overview of the risks associated with data quality issues, and how that risk is introduced across the enterprise.

Managing Scorecard Views

Essentially, each of these views composing a data quality scorecard require the construction and management of a hierarchy of metrics related to various levels of accountability for support the organization's business objectives. But no matter which scheme is employed, each is supported by describing, defining, and managing base-level and complex metrics such that:

- Scorecards reflecting business relevance are driven by a hierarchical rollup of metrics.
- The definition of metrics is separated from its contextual use, thereby allowing the same measurement to be used in different contexts with different acceptability thresholds and weights.
- The appropriate level of presentation can be materialized based on the level of detail expected for the data consumer's specific data governance role and accountability.

SUMMARY

Scorecards are effective management tools when they can summarize important organizational knowledge as well as alerting the appropriate staff members when diagnostic or remedial actions need to be taken. Part of an information risk management program would incorporate a data quality scorecard that supports an organizational data governance program; this program is based on defining metrics within a business context that correlate the metric score to acceptable levels of business performance. This means that the metrics should reflect the business processes' (and applications') dependence on acceptable data, and that the data quality rules being observed and monitored as part of the governance program are aligned with the achievement of business goals.

These processes simplify the approach to evaluating risks to achievement of business objectives, how those risks are associated with poor data quality and how one can define metrics that capture data quality expectations and acceptability thresholds. The impact taxonomy can be used to narrow the scope of describing the business impacts, while the dimensions of data quality guide the analyst in defining quantifiable measures that can be correlated to business impacts. Applying these processes will result in a set of metrics that can be combined into different scorecard schemes that effectively address senior-level manager, operational manager, and data steward responsibilities to monitor information risk as well as support organizational data governance.

NOTES

1. U.S. Dept. of Defense, "DoD Guidelines on Data Quality Management," 1997, accessible via www.tricare.mil/ocfo/_docs/DoDGuidelinesOnDataQualityManagement.pdf.
2. "2006 ACFE Report to the Nation on Occupational Fraud and Abuse," www.acfe.com/documents/2006-rttn.pdf.

3. Herbert, Brian, "Data Quality Management—A Key to Operator Profitability," *Billing and OSS World,* March 2006, accessible at www.billingworld.com/articles/feature/Data-Quality-Management-A-Key-to-Operator.html.
4. Inserro, Richard J., "Credit Risk Data Challenges Underlying the New Basel Capital Accord," *RMA Journal,* April 2002, accessible at www.pwc.com/tr/eng/about/svcs/abas/frm/creditrisk/articles/pwc_baselcreditdata-rma.pdf.
5. Ernst & Young, "Raising the Bar on Catastrophe Data," 2008, accessible via www.ey.com/Global/assets.nsf/US/Actuarial_Raising_the_bar_catastrophe_data/$file/Actuarial_Raising_the_bar_catastrophe_data.pdf.

CHAPTER 4

Total Quality Management Using Lean Six Sigma

Praveen Gupta

INTRODUCTION

Total Quality Management (TQM) has been defined as management of activities, results, and decisions for quality throughout the organization. TQM in the financial industry would mean managing all aspects of finance business to achieve business objectives, including profitable growth. The financial industry has an edge over other industries, such as manufacturing, where due to the nature of the business, the error rates are much lower (about 0.05, compared to about 0.2 in manufacturing). Any minuscule mistake can result in a huge adverse financial impact. That is why to a great extent the financial industry is regulated through a variety of checks and balances. Before one deploys TQM in the financial industry, one must first understand the definition of quality in the financial industry.

Quality means different things to different stakeholders. The most important aspects of the financial industry are managing risks and accuracy of operations. Thus, from the customers' perspective, quality could be defined as consistency of accurate information and reporting. From the stockholders' perspective, operations should be virtually risk free. The quality goals may appear to be difficult to achieve however striving for them is definitely possible. For employees quality would mean minimizing their operational glitches and errors, and for suppliers or partners quality would mean dependability of the business relationship through clarity of expectations, transparency, and measurable verification. One can see that quality means different things to different stakeholders and is defined for various stakeholders and at various stages of the operations.

The financial industry has been practicing quality by complying with the regulatory requirements. Quality through compliance to the requirements helps but does not necessarily ensure best performance. It does not question strategic intent of activities. Research shows that until recently the financial sector did quite well in terms of net profitability. Today, the financial sector is considered to be volatile and tainted with risks. The recent mortgage crisis is drawing more attention to the financial sector. The troubling segments within the financial sector, savings and loans, mortgage investments, and real estate investment trusts (REITs), have raised awareness for the much-needed quality management in the entire financial sector.

According to the Federal Deposit Insurance Corporation (FDIC), 34 banks have closed since 2000. However, this list does not include catastrophic failures of large institutions such as Bear Stearns and Countrywide Bank in 2008. The failure of financial institutions is not entirely due to their internal actions. Instead, it is the result of the interaction among various institutions, poor controls, and lack of quality assurance.

Failures of large institutions such as Bear Stearns, and Countrywide make us believe that failure was caused by inconsistency in following internal procedures, having sufficient internal controls to prevent continuation of malpractices, and ignoring key performance indicators. In other words, various processes and functions were not meeting their quality expectations. Interestingly, a few weeks before Countrywide was acquired by Bank of America, it advertised a job for vice president of continuous improvement, implying that financial institutions do need quality practices in order to prevent unintended and unmonitored activities and their risks. It was too late by then!

The financial industry is a transaction-driven industry where many events take place very quickly and require virtual perfection. We cannot afford to have errors in percentage points. Financial institutions do deploy business processes, utilize information to deal with customers, and require discipline to execute decisions effectively and accurately. TQM addresses quality of activities through process management, quality of results through performance measures, and quality of decisions through commitment to continual improvement using a variety of quality management tools.

A TQM initiative in the financial sector must include understanding of the process management for achieving excellence and continual improvement through Six Sigma and Lean-type methodology, and performance measures through service scorecards that will provide a high-level picture to the executives. The following tools will facilitate the implementation of TQM.

PERFORMANCE TARGETS

Conventionally, TQM implied managing the process to deliver acceptable process output or the product. That would be ensured through quality inspection, control, and assurance techniques. TQM meant planning the output, doing the activities, and controlling the output within established specification limits. Such a model of TQM worked well in the absence of true global competition, where customer expectations were moderate. In terms of process yields, they could be practically around 95 percent. Due to excessive verification activities, appraisal cost went up, resulting in a high cost of quality. Process expectations were more driven by the process capability rather than the customer's needs.

However, today, customer expectations are much higher, and in order to achieve excellence, one must establish clear targets. The targets are driven by the customer expectations. In order to learn customer expectations, Kano's model has been a powerful method where customer demands are classified into three categories: assumed, spoken, and desired. One can think of these three types of customer requirements as minimal, paid for, and wished for. It has been learned that only when customers get what they wished for, in addition to the minimal and paid for, they love suppliers' performance.

Noriaki Kano modeled the relationship between customer satisfaction and customer requirements. Accordingly, most of the customer requirements are unspoken. What we are told is little, and the customers expect a lot more than what they ask for. Kano's model provides an excellent platform to achieve the organizational objective of having its customers patronize its services or products, and as a result lead the organization into becoming a best-in-class service provider. According to the Kano's model, customers have the following three types of requirements:

1. *Unspoken assumed minimal requirements.* When a customer seeks financial services, there are certain assumptions that the service will be timely and accurate, and will not cost excessively or cause loss of his wealth. Knowing the risks associated with financial services, if the service provider does not have credibility and capability, the customer would not want to work with the service provider. These unspoken assumed requirements are called "dissatisfiers."
2. *Spoken and paid-for marketplace requirements.* The marketplace requirements are commonly known expectations built through branding or general awareness of the industry. Customers know that these days there are options for financial services. For example, if they go to Charles Schwab's web site or E*TRADE for services, they have learned from the web site and its advertisements, and have been told what to expect. Customers pay for the promised services and expect results. In the absence of promised or expected returns or results, customers feel dissatisfied. Educating customers about risks, setting right expectations, and proactively following up with customers may help in fulfilling spoken customer requirements.
3. *Unspoken wished-for or love-to-have requirements.* In the age of customer and supplier relationship management, organizations are learning to love their business partners. Any relationship requires knowing what your partner loves to have; similarly, in the case of customer relationship management (CRM), the project team must learn what it is that customer would love to have. Also, as a financial service provider, one must learn about love-to-have requirements of internal or external customers. The love-to-have requirements may be in terms of proactive communication, extra income, convenience of information, on-call service, advance risk mitigation notices, or a surprise gift.

Exhibit 4.1 shows Kano's model of customer requirements. The x-axis shows level of effort by the service provider, and the y-axis represents extent of customer satisfaction. The intersection of two axes represents "do not care" on the y-axis. One can see that by providing the services to meet the assumed requirements, the best one can do is to achieve customer's ignorance of performance. As to the spoken requirements, in terms of getting financial service or output from a preceding process, customer satisfaction grows proportionately. In other words, the more we satisfy the customer's stated requirements needs, the more they are satisfied with the services. The final element of the customer requirements, "love to have," is beyond the spoken requirements. Customer satisfaction exponentially grows with the provision of "love-to-have" requirements. Customers love surprises, brag about the service provider, and bring in new business or goodwill through word of mouth.

All three requirements must determine the performance targets for designing the financial service and the process to deliver it.

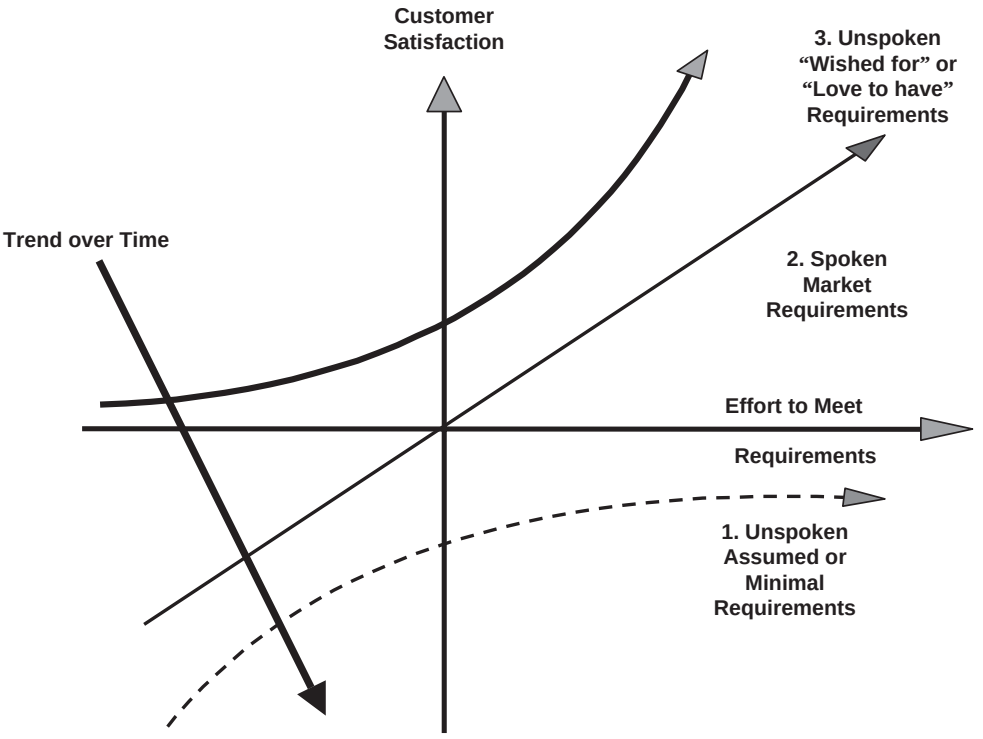


EXHIBIT 4.1 Kano's Model

PROCESS FOR EXCELLENCE

Understanding excellence is critical before designing a process for achieving excellence. Excellence does not imply "zero" errors, as people can achieve them with excessive verification. A "zero" defect process may be functional but sloppy at best. Thus, excellence must be understood as perfection that means being on target. Once the targets are defined as per the aforementioned Kano model, the process must be designed to achieve the target performance. The 4P model of process management developed in 2006 is a great way to achieve excellence.

The 4P model (prepare, perform, perfect, and progress) offers a better implementation of process thinking than the typical plan, do, check, and act cycle. Exhibit 4.2 shows the 4P model representing aspects of process management for achieving excellence. *Prepare* implies doing homework or setting up a process to achieve target performance. For example, if a process to distribute dividends must be designed, preparation must include getting all the information required; developing a system for scheduling, printing, enveloping, and mailing; an error-free and streamlined process flow; and defining skills and identifying the right personnel to perform. Preparation is a critical element of managing a process for excellence. Without good preparation, errors occur and target performance is missed. *Perform* implies doing things well, instead of just doing it. During the *perform* aspect, critical process steps must be identified, measured, and monitored against specified target values. *Perfect*

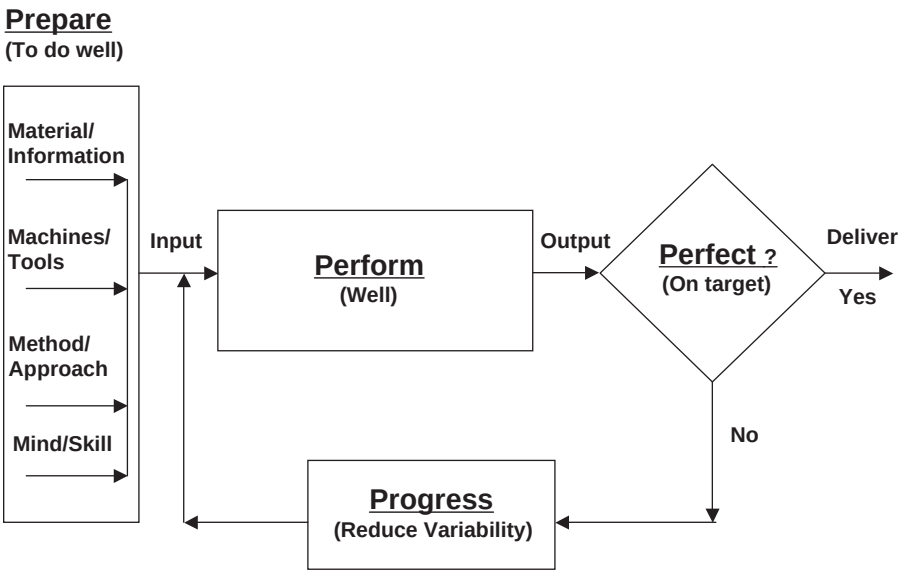


EXHIBIT 4.2 4P Model for Process Excellence

represents evaluating performance against the specified targets. Initially, aiming for a target performance may appear to be a difficult task; however, the process must be designed through inputs (good preparation) and activities (perform) such that the output lands at or close to the target value (perfect). Close to target performance keeps the process output away from the lower and upper specification limits, resulting in lower failures and thus the cost of failures.

One of main benefits of implementing the 4P model is the attitude changing from "acceptability" to "excellence." The excellent process output will lead to better profit margins than the acceptable process output due to reduced cost of appraisals.

The 4P model has also been proven to be helpful in identifying measures of effectiveness at the process level and system level. The measures of effectiveness can be established at the input, in-process, or output stages of a process. For example, for the mortgage approval process, the measures of effectiveness could be financial strength of the borrower, timely verification of borrower's records at the input level, compliance to establish activities or regulatory requirements prior to approving the mortgage at the process level, and payment schedule compliance at the output level.

PROCESS IMPROVEMENT

Once a process is designed to achieve excellence or the target performance, variation may occur and the performance may be affected. Six Sigma and Lean are two known methodologies for improving processes. The Six Sigma methodology was developed to achieve virtual perfection—that is, close to the established performance targets—and Lean is a variation of the Toyota Production System or Henry Ford's Assembly Process. The Toyota Production System is to achieve perfection as

well through process designs without waste of time, material, equipment, human resources, or space.

Six Sigma

According to the early documents at Motorola, where Six Sigma was first used, a simpler definition stated, "Six Sigma is our Five Year Goal to approach the Standard of Zero Defects, and be best-in-class in EVERYTHING we do." Accordingly, we can define Six Sigma as *an approach to achieve virtual perfection fast, and be the best in class in everything*. A tactical definition based on statistical analysis, Six Sigma can be defined as having the process capability twice as good as required.

At Motorola, Six Sigma was originally defined as a measure of the goodness of products and services. Higher sigma means better quality of a product or service, and lower sigma means poor quality of a product or service. The original Six Sigma initiative included leadership drive, the Six Steps to Six Sigma methodology, and related measurements. The six steps are:

1. Define your products or services.
2. Identify your customers and their critical needs.
3. Identify your needs and resources.
4. Map your process.
5. Remove non-value-added activities and use error-free methods.
6. Measure the sigma level, and continue to improve the process if the sigma level is less than 6.

The statistical definition focuses on tactics and tools, while the original definition focuses on the intent and methodology of Six Sigma. The intent of Six Sigma is to achieve a significant improvement fast by using the commonsensical DMAIC (define, measure, analyze, improve, and control) methodology. Critical success factors for deploying the Six Sigma methodology include:

- Passionate commitment to Six Sigma.
- Common language to be used throughout the organization.
- Aggressive improvement goals that will force continual process reengineering.
- Innovation, not the statistics, as the key to achieving dramatic improvement.
- Correct metrics for assessing the next steps to achieve dramatic results.
- Communication to maintain continuity and interest in the Six Sigma initiative.

DMAIC Methodology

DMAIC is a five-phase improvement methodology. Experience shows that the *define* phase is the essential phase for achieving dramatic improvement quickly, and the *control* phase is the most critical phase for realizing return on investment.

The success of the DMAIC methodology depends on working well on the right projects. The right project is the one that can result in a significant return on investment. Thus, the first priority is to identify the right projects to work on that will have an impact on the bottom line and generate savings for the business. Several potential projects must be identified and evaluated based on a cost and

benefit analysis. A simple measure, such as the project prioritization index (PPI), can be used to prioritize projects according to the following equation:

$$\text{PPI} = (\text{Benefits/cost}) \times (\text{Probability of success/Time to complete the project in years})$$

At a minimum, the PPI should exceed 2 to ensure a return on investment. Initially, one can find many projects with PPI greater than 4, thus making it somewhat easier to realize savings.

Once the project is selected, the team representing various functions is formed to work on it. The team receives Six Sigma training at the Green Belt level while working on the selected project. During the *define* phase, the team develops a clear definition of the project, the project's scope, the process map, customer requirements, SIPOC (suppliers, inputs, process, outputs, customers), and a project plan. In other words, in the *define* phase, customer requirements are delineated and a process baseline is established.

In the *measure* phase, we establish the sources of information, the performance baseline, and the opportunity's impact in terms of cost of quality. The performance baseline is established in terms of first-pass yield (FPY), defects per unit (DPU), defects per million opportunities (DPMO), sigma level, and basic statistics such as mean and range or standard deviation.

In the *analyze* phase, the focus is to examine patterns, trends, and correlations between the process output and its inputs. A cross-functional team performs the cause-and-effect analysis using the fishbone diagram. The purpose is to identify the root cause of the problem and necessary remedial actions to capitalize the opportunity. At the end of the *analyze* phase, the team is able to establish a relation such as $Y_{output} = f(x_{inputs})$.

While analyzing data, one should look into whether the excessive variation or inconsistency is normal in the process or has temporarily crept into the process. If the inconsistency is normal, a thorough capability study is required, and perhaps the process needs to be redesigned. If the inconsistency is exceptional, the process will need adjustment. Failure Mode and Effects Analysis (FMEA) is also used in the *analyze* phase (or subsequent phases) to anticipate potential problems or risks, as well as to develop actions to reduce risks of failures.

The first three phases of the DMAIC methodology help in gaining a better understanding of the process and learning the cause-and-effect relationship between the output and input variables. The *improve* phase enables the development of alternate solutions to achieve the desired process outcomes.

Typically in a non-sigma environment, we jump to solving the problem directly without defining and understanding the process well. Without such an in-depth knowledge of the process, solving a problem becomes a game of luck. Experimenting techniques are used to fine-tune the relationships or optimize the process recipe. However, such experiments are rarely required if nonstatistical tools have been effectively utilized in the early phases.

The *control* phase is employed to sustain the improvement utilizing effective documentation, training, process management, and process control techniques. In the *control* phase, a score of the process or business performance must be maintained, and the sigma level must be continually monitored. The *control* phase is also an

EXHIBIT 4.3 Key DMAIC Tools

Phase	Tools
Define	Pareto, process map, Kano's analysis, SIPOC, CTQ
Measure	DPU, DPMO, Sigma level
Analyze	Root cause analysis, FMEA, scatter plot, visual correlation
Improve	Comparative and full factorial experiments
Control (Sustain)	Process thinking (4P model), review, control charts, scorecard

opportunity to engage senior management in the Six Sigma journey for support and aggressive goal setting for identifying further opportunities for improvement.

The DMAIC methodology incorporates numerous tools. Exhibit 4.3 summarizes simple yet powerful tools in the DMAIC methodology.

Besides tools, there are three measurements uniquely identified with the Six Sigma methodology: DPU (defects/errors per unit), DPMO (defects per million opportunities), and sigma level. The DPU is a unit or the output-level measurement, DPMO is the process-level measurement, and sigma is a business level measurement. Sigma provides a common theme for the organization and requires a lot of improvement to show a positive change. The customer cares for DPU, the process engineer needs to know DPMO, and the business executives drive the sigma level. All of these measurements can be used to communicate performance expectations and progress throughout an organization.

The most commonly used measurement driving improvement in an organization must be DPU. The DPU is converted into DPMO based on the process or product complexity, and the DPMO is transformed into the sigma level for establishing a common performance measurement across all functions in an organization. Key DPMO associated with the sigma levels are 66807 for Three Sigma, 6210 for Four Sigma, 233 for Five Sigma, and 3.4 for Six Sigma.

Lean

Lean thinking has been practiced in U.S. manufacturing since the 1980s, and since the 1960s in Japan. Lean-like principles were first deployed by Ford while standardizing parts production and assembly operations in 1910s and 1920s. In the United States, Lean was initially known as Just-in-Time (JIT) manufacturing, which was successfully implemented in distribution of parts by delivering customer-ordered parts when and where needed. When implementing JIT principles, the focus shifted from producing to a forecast to producing to the customer order. This thinking was also called a pull system (JIT) versus a push system (forecast). One can see that in the financial sector operations by nature are forecast driven. However, some Lean tools are still applicable for reducing non-value-added activities.

Lean is intended for setting up waste-free operations, whether manufacturing or finance operations. Waste-free operations means providing what is needed when it is needed and maintaining a rhythm at a given throughput level in response to customer demand, rather than using maximum capacity to stay busy. Thus, one of the objectives of Lean implementation is to design a system that can be in rhythm with

EXHIBIT 4.4 Comparison of Lean and Six Sigma

Lean	Six Sigma
Implemented for efficiency and reduction in wasteful activities	Implemented for effectiveness and reduces waste in an activity
Driven by middle management	Driven by leadership
Supports the target of achieving virtual perfection	Provides targets for virtual perfection
Synchronizes the resources utilization	Synchronizes skills with the customer service improvement
Requires personal commitment to challenge current processes	Requires passionate and inspirational commitment from CEO to create mind for virtual perfection
Impacts selected processes for speed and value	Impacts all aspects of business products and processes

customer demand. Rhythm implies minimal wait time or interruptions in operations. Lean minimizes changes, abnormalities, or fluctuations in the flow of material or information as well as use of wrong tools or models; ensures visibility of operations and deviations; is an immediate remedy to unacceptable activities or outcomes; and emphasizes planned and leveled workload.

Exhibit 4.4 lists comparative aspects of Lean and Six Sigma. For example, the management thinking for Lean implies speed and flow, while for Six Sigma it is quality and time. If we combine Lean and Six Sigma, one can think in terms of quality and flow, which will minimize time and speed up the process. If one were implementing Lean alone, the solution could lead to very fast process increasing risks of failure, while implementing Six Sigma alone may result in virtually no risk but it may take forever. Thus, a combination of Lean and Six Sigma will allow risk and speed optimization.

Role of Innovation

Most process improvement or TQM activities are designed to achieve a sustainable process for achieving excellence. However, in the dynamic financial industry and changing global economy, one must be able to change well-established processes or products quickly to meet customer needs. This requires incorporation of innovative thinking in practicing quality principles. Innovative thinking entails creativity and deployment to maximize return on investment in improvement. Innovation has been considered an adversary to improvement. Once again, we must learn to balance consistency and creativity, as well as improvement and innovation. Both are simultaneously required to sustain excellence and achieve best-in-class performance.

SUMMARY

Conventionally, TQM has been understood as a group of quality activities without a clear target and deployed for years before one could see results. However, in

today's Internet age, solutions are demanded quickly. Therefore, TQM must evolve to include defining performance targets, designing processes to deliver target performance, and improving suboptimal processes quickly using powerful methodologies like Six Sigma and Lean with innovative thinking. In order to earn customer loyalty and grow business, innovation enables us to make our solutions a distinct competitive advantage. TQM must incorporate innovation for developing breakthrough designs and improvement results.

CHAPTER 5

Reducing Risk to Financial Operations through Information Technology and Infrastructure Risk Management

Brian Barnier and Richard Marti

INTRODUCTION

The risks to a business related to its information technology (IT) and related physical infrastructure have never been greater as increased automation, IT complexity, globalization, and more tightly linked partners all make it crucial that the IT "factory" of a financial institution run smoothly and be resilient enough to take advantage of the next business opportunity. The proactive IT leader faces challenges to effectively manage IT and infrastructure risk, from the basic "What does 'IT risk' mean?" to the complex "How do I simplify risk management when the business itself is so complex?"

While the challenges are nontrivial, the risk-aware IT leader has never had such opportunity to drive change. Headline-grabbing examples of business outages related to IT and infrastructure, business needs for expansion, and compliance pressures all demand action. While these challenges have never been greater, the help available is also better than ever. Automation and more streamlined risk management approaches make significant leaps possible to bring improved IT return with less risk to the business.

This chapter seeks to help you accelerate improvements in IT and infrastructure management by drawing on both recent trends in IT and decades of broader experience in risk management in industries ranging from manufacturing to hospitals.

THE PROBLEM

In talking about risk, let's start with the basics—risk is the probability that something will happen multiplied by the impact if it occurs. It's important to note that a "risk" is really a chain. This chain has been described differently by various authorities, but a good example is an actor, a threat, timing, impact, and resulting damage. With this understanding of a chain, we can see that using terms like *disaster recovery risk, hacking risk,* and *reputation risk* are not quite accurate. Hacking is a threat. Disaster

recovery is a response to an impact. Reputation is damage. To help you communicate more clearly in your organization, just think about the chain.

With this as background, we can consider risk in two senses:

1. In the financial management sense, as one half of the risk-and-return equation that lies behind all business decisions and objectives. The enterprise is in business to achieve objectives like revenue, profit, growth rate, share price, market share, brand equity, and customer satisfaction. This is very similar to the way you might manage your retirement plan, accepting more return relative to risk.
2. In the operations management sense as problems that may arise in producing or distributing quality product. This is similar to the way you might follow a planned maintenance schedule on your car.

In thinking about it in these two ways, we can consider how it is viewed by various roles in the organization:

- At the chief financial officer (CFO) level, the need is to help the business take greater advantage of market trends in the effort to achieve the objectives.
- At the chief operating officer (COO) level, the need is to be able to continue to deliver product and/or service to customers (and partners) in the face of a range of threats to the IT and physical infrastructure of the business. (Threats to people are also a concern but are outside the scope of this chapter.) This includes compliance with laws and regulations, industry standards, and contractual requirements.
- At the sales executive level, the need is to reliably step up to customer contract requirements to be able to win and expand business.

In short, risk is taken in pursuit of return. However, that risk must be actively managed; otherwise, the return will not be achieved and loss will be more likely.

As enterprises navigate risk-return waters, they interact with industry trends and issues—to avoid being overwhelmed by them (e.g., competitors merging or new government requirements) or take advantage of them (e.g., geographic or product expansion).

While some trends are truly specific to an industry, others are expressed differently in a given industry but really have much in common across industries—especially when it comes to the financial and operational impacts. These include acquisitions, consolidation/cost cutting, globalization, automation, integration, and compliance. These and related trends drive a heightened awareness of risk because they all involve significant change in an enterprise. Change can bring good or bad outcomes. Because such initiatives introduce so much change, they scream for careful risk management. Change that happens in compressed time frames causes greater risk. Change as part of cost cutting is still greater risk as knowledge disappears and both business and IT processes are disrupted. A risk approach is used to analyze and respond to such activities to maximize the potential for positive outcomes. Again, risk equals probability of something's occurring multiplied by the business impact if it occurs.

So let's consider each of these changes in view of the risks *to* IT and related physical infrastructure *and* the risks *from* IT and infrastructure to the business change initiative.

Acquisition

An acquisition places demands on IT to integrate systems, consolidate some business functions into a single system, expand other systems to meet new feature requirements and transaction volumes, and more. This must all be done in a time-compressed environment to meet announced commitments to shareholders and others.

IT and infrastructure can put timely acquisition integration in jeopardy if it lacks the resilience—operational stability, availability, protection, and recoverability needed to meet the business requirements.

Regulatory compliance adds complexity in merging inconsistent policies and procedures across the organization, and new regulation due to new product types, geographic areas of operations, and other factors. These problems are usually complicated by lack of expertise and especially extra expertise during the acquisition integration. Whether for contracts, industry standards, or regulations, a firm does not want to be out of compliance during an acquisition.

Consolidation

Consolidations and cost cutting driven by postacquisition pressures, business contraction, or rationalization can introduce a range of risks from project-oriented problems (e.g., causing operational instability) to postconsolidation impacts (e.g., inadvertently closing a more resilient facility and leaving open a less protected one in pursuit of short-term cost savings).

Consolidation requires operational change management on steroids. Lacking solid IT controls and processes, much unneeded risk will be injected into the business—especially under pressure to "cut cost fast." In these cases, the business owners must be made fully aware of what knowledge will be lost, process disrupted, or systems made vulnerable due to resource cuts. Further, cost-cutting-driven risks must be made visible to leaders in sales, partnering, and supply chain roles. They must understand not only the risks of cost cutting, but impacts on existing or potential contracts as well as industry standards.

This is not to say that all cost cutting will drive up risk. To the contrary, *carefully* applied risk management can actually help streamline process, reduce complexity, reduce waste, and save time. This is a legacy of the quality improvement side of operations risk management. Just don't skip the "carefully" part.

Expansion

Expansions are fast-paced by nature in order to take advantage of some opportunity. This can be geographic or product oriented. From the business aspect, they can open new national government or liability requirements that place new needs on IT.

IT systems are pressed for resilience in expansion. While some new products have ground-up new IT, more often current capabilities are stretched to cover the new requirements. Operational stability and availability are often stressed by new

transaction volumes. Data may have new protection requirements due to regulation or third-party needs.

Lack of IT strategic planning is one of the major challenges for expansion.

Globalization

Globalization expands on the new product expansion to respond to industry patterns in general—competitor actions, supply chains, new notions of extended enterprise, or changing end-customer patterns.

IT and physical infrastructure is stressed in this environment to maintain operational stability, protection, and recoverability in more distant environments, with higher support costs, less stable local resources, and other hazards.

Globalization has also introduced a new dimension to information security. Data breaches are an ongoing concern due to lack of sufficient control visibility in emerging countries, especially when facilities and employees are growing quickly.

Automation

Automation reflects the reality that enterprises are simply using more technology to improve quality and reduce cost. In doing this, business process is becoming more dependent on the underlying IT and physical infrastructure.

Many end users only see the dependency on IT as far as the application software. However, the real dependency is an entire technology "stack" including middleware, servers, storage, networking, buildings, energy, and such. Everything must work for the business process to generate revenue. Example problems include failure to include IT service management software on the critical IT systems' list and errors in applying new virtualization.

Here, automation can accomplish two things: (1) save millions of dollars on external auditor fees and reduce human errors—if done right, and (2) be used to "build in" risk management into those underlying IT and physical infrastructure systems to more efficiently monitor and remediate threats.

Integration

Integration comes along to cross the islands of automation. Without this, there are weak links in the business process chain. Integration improves quality, speed, reliability, and business reach (internally or across enterprises in partnerships).

For IT and physical infrastructure, integration takes the concept of a technology stack under a business application to a far broader level. Now multiple stacks must work together (sometimes with common supporting elements). If you now have a picture in your mind of a large industrial factory or an integrated web business, you have the right idea. From a risk perspective, the failure analysis is now quite complicated. It becomes more difficult to track the impact of a potential threat to one asset or other aspects of the system.

Compliance

Compliance takes the trends discussed earlier to an added level of monitoring and reporting complexity. From the business perspective, they can impose an absolute

bar to revenue as well as cause penalties or reputation damage from failing to comply with contract requirements to customers or regulatory requirements.

Technology Changes

Reorganizing technology operations management is a popular topic—shared services, data center consolidation, green, and more. These have great opportunity, but also significant risk in the following three stages:

1. The initiative is vulnerable to design risk. Does it capture the right requirements? Does the solution actually address the requirements?
2. Risk from implementation problems due to either project management or operational errors before the change is fully vetted and stable.
3. Postimplementation from problems that arise during maintenance or from interaction with other systems.

As organizations seek business advantage through technology, they seek resilience and cost reduction through technology platform improvements. Examples include virtualization, mobility, web 2.0, cloud computing, service-oriented architecture (SOA), collaboration, and more.

Each of these platform changes has the potential to bring business value and reduce risk. In addition to these points, two more considerations are:

1. *Cost and risk management are disconnected.* The classic example is consolidation for cost reduction that leaves multiple single points of failure.
2. *Failure to look at the entire technology stack supporting an application when one element is changed.* For example, an SOA initiative can underachieve results when the focus is only on resilience and flexibility at the application layer, without looking at the resilience in the rest of the IT and physical infrastructure stack.

In all of these cases, problems can appear in two ways—process and outcome. For example, to prevent disease transmission in hospitals, risk managers look at both failures of proper hand washing and rates of diseases acquired during patient stay.

Process Problems

Process problems are like failures to properly wash hands. It can also be the lack of a process to identify poor hand washing as a problem. In an IT sense, examples include:

- Lack of industry standard IT control framework (e.g., Information Technology Infrastructure Library [ITIL] or Control Objectives for Information and related Technology [COBIT]).
- Lack of clear internal definition of IT and infrastructure risk, often reflected in IT silos with limited views of threats.
- Insufficient learning from past external or internal failures.
- Failures repeated too often—operation instability.
- Cannot get ahead of weaknesses flagged by IT auditors/regulators.

- Missing types of threats—losses are too often surprises.
- Missing preventive and corrective controls, especially on creeping threats.
- Difficulty communicating across IT areas or with business areas.
- Difficulty communicating with partners to avoid passing failures through integrated systems.
- Lack of training.

Outcome Problems

The following outcome problems are reflected in measures like increased rates of disease transmission:

- Not detecting threats until too late.
- Losses in penalties and fines.
- Revenue lost from failure to meet customer needs.
- Unable to expand or move quickly on opportunities.
- Reputation damage.

These problems jump to the organization's attention in three ways:

1. *Through an audit or test*. You're embarrassed, but no real harm (unless there is a regulatory fine involved).
2. *Through a near miss*. You started to fail, but heroics saved the day. What will happen tomorrow?
3. *Through an actual incident*. Depending on how quickly it is detected and resolved, the damage in terms of costs, fines, penalties, lost revenue, or reputation can vary considerably.

Wouldn't it be nice to have visibility into potential problems before they occur?

RISK SOURCE AND ROOT CAUSE

With all this said about activities that increase risk and ways to observe risk, it's helpful to consider how to characterize, classify, or report risk in a way that is more actionable.

The following principles might be helpful:

- *Separate cause from effects*. For example, reputation damage is a result of a risk being realized, not a source.
- *Separate proximate (or nearest) cause from root cause*. For example, a penalty is incurred because a replacement part was not delivered to a customer on time. This could be classified as "replacement parts process failure." However, with a little more examination, we see that an order-entry IT system failed. Still more looking shows a transaction processing failure. Additional analysis shows a transaction volume spike that overwhelmed software and/or hardware. Taken further, it could be blamed on "too many customers want to buy our product." Yet taking it that far would be outside of your actionable space. So you would

want to class the threat source as transaction volume. Then work a remediation plan to address that in an end-to-end business context (otherwise, you might fix the servers and then the software would fail).

- *Separate actual threat actors and actions from the people, processes, and things against which they act.* For example, you might hear someone talk about "network risk" or "server risk." These might be a helpful way to aggregate all the risks faced by the server administrator. However, it does not get people thinking about the range of threats and how to predict, detect, and correct those incidents. You can have categories of things (sometimes termed *assets* or *resources*) impacted by a threat, but you also need to look at threat sources.
- *Focus on actual operations, not just loss byproducts.* For example, a beer brewer does not just analyze customer complaints; they also evaluate quality controls at multiple steps in the brewery and back into the supply chain. An automobile insurer does not only look at claims losses; they also are active in car, road, and driver testing.

Taking these together will make it easier to focus on *real threats to real operations*, conduct *meaningful scenario analysis,* or evaluate *interdependencies* among business activities and related threats.

This is essential to providing a business context of the problem that is compelling to a business line owner, corporate risk manager, and various IT leaders.

RISK MANAGEMENT

Awareness, culture, organization, governance, techniques, and tools all play a role in managing IT risk. A comprehensive solution must address all these areas to be successful.

Awareness/Culture

One of the major challenges in the risk management framework is awareness and/or culture gap. Globalization and lack of training and skills have widened this gap. External auditors view this as a significant deficiency while evaluating enterprise risk posture for publicly traded companies. We strongly recommend addressing this real issue as a part of the enterprise risk management (ERM) effort.

Organization/Governance

The most effective risk management program should include a top-down risk-based approach that involves the board of directors and senior management. Sound integrity and ethical values, particularly of top management, are developed and understood and set the standard of conduct for external reporting. The board of directors should actively evaluate and monitor risk of management override of internal control. Management should establish triggers for reassessment of risks as changes occur that may impact company objectives. Most of all, the governance process should become more "risk aware" and be informed of both risk and return implications of investment decisions and operational control.

Techniques and Tools

Risk management for IT and related physical infrastructure continues to develop as a discipline. In doing so, it builds on decades of risk management in areas like industrial operations, finance, insurance, and others. By borrowing from the various business understandings of risk management, it gains both a knowledge base and ease of communication with business areas.

As this is a large subject, a few examples are mentioned here along with a table of common standards and practices that follows.

Risk management quality control will frequently employ Six Sigma and other systematic approaches commonly found in industrial settings. These are making headway in financial institutions as well. The Six Sigma body of knowledge, originally developed at Motorola, provides helpful approaches for making decisions in view of risk, problem diagnosis, root cause understanding, and evaluating different aspects of a system for potential risk and failures. While an aspect of the Six Sigma approach involves detailed statistical analysis (from which the technique is named), much of the Six Sigma approach is simply helpful ways to understand dependencies and thus risk within a system. This is basic knowledge that everyone needs to reduce risk and improve quality of service in any environment.

To introduce you to Six Sigma land, here are some of the key concepts:

- For ongoing operations, DMAIC is a hallmark approach of Six Sigma. It stands for define, measure, analyze, improve, and control. Another aspect of Six Sigma focuses on new projects and emphasizes *design* and *verify*.
- Six Sigma is highly process oriented. The acronym for this is SIPOC—supplier, input, process, output, and customer. The SIPOC process analysis approach includes a kit of techniques for root cause analysis that are highly applicable to IT and infrastructure risk. These include the fishbone diagram, cause-and-effect analysis, and Failure Mode and Effects Analysis (FMEA).
- FMEA is similar to traditional risk analysis (likelihood $\times$ impact) but adds the additional factor of delectability. This is an important insight. Considering this factor in risk rankings increases the priority of attention to those threat conditions that can sneak up on you to hurt your operations.

Within IT, there are special challenges. These start with the need to bring together various IT silos such as change management, access security, disaster recovery, network protection, availability, and more. While industrial facilities have long been training to take an end-to-end, assembly-line view of threats to production or distribution, IT has lived in its silos. To help bring IT risk areas together, COBIT is the most widely used industry standard since its introduction over a decade ago.

As described in the executive overview of COBIT 4.1:

> *For many enterprises, information and the technology that supports it represent their most valuable, but often least understood, assets. Successful enterprises recognize the benefits of information technology and use it to drive their stakeholders' value. These enterprises also understand and manage*

> *the associated risks, such as increasing regulatory compliance and critical dependence of many business processes on information technology (IT).*
>
> *The need for assurance about the value of IT, the management of IT-related risks and increased requirements for control over information are now understood as key elements of enterprise governance. Value, risk and control constitute the core of IT governance.*

Within the IT Governance Institute's (ITGI's) approach to enterprise governance of IT, risk management is one of the five focus areas. Risk is addressed in both the financial sense of risk and return in investment portfolios and in the operational sense of risk to daily execution. With such an emphasis, the ITGI released a new IT risk governance and management framework that addresses both the strategic and practical issues in IT risk management.

This contribution is intended to be a more robust bridging of the gap between general business risk management approaches and those for various IT-related areas. It includes a risk management process that provides the missing link between ERM and IT management and control, fitting into the overall IT governance framework approach of ITGI that is built on COBIT and Val IT. It addresses the full risk life cycle and seeks to make it relevant to risk managers, business process owners, CFOs, IT operations leaders, and auditors. Following the style of other ITGI frameworks, sections include risk taxonomy, domains and processes, RACI (responsible, accountable, consulted, informed) charts, maturity models, and supporting implementation appendices. With this level of content and connectedness, it is the new benchmark for IT-related risk governance and management.

Two very useful features of the new guidance are (1) the pains it takes to bring clarity to often confusing usage of risk terminology and (2) its guidance about what "good" looks like.

COBIT also helps simplify communication both inside and outside the enterprise. First, it has been mapped with other IT management techniques such as ITIL for IT service management. Second, it can link with business risk management approaches such as COSO ERM (Committee of Sponsoring Organizations of the Treadway Commission enterprise risk management) from the United States, A Risk Management Standard (ARMS) from the United Kingdom, or AS/NZS 4360 from Australia and New Zealand. A mapping document connects COBIT and COSO ERM. Third, because COBIT is also used by IT auditors, it also helps simplify internal compliance. Fourth, because COBIT is used so widely, it can help bring together partners, customers, and suppliers in the extended enterprise.

As to other standards and practices, Exhibit 5.1 provides a high-level matrix that describes some of the frameworks that are being used in many organizations.

CLOSING COMMENTS

In closing, it has been observed on many occasions that effective risk management is a daunting challenge due to the complexity of an enterprise's organization structure, geographic distribution, and business processes being evaluated.

When managing the IT and related physical infrastructure risks on which those processes depend, the complexity grows with the complexity of that infrastructure. Many firms approach any change in IT with trembling for fear of unintended damage.

Despite this, an IT risk leader can quickly drive value in his or her organization by keeping a few guideposts in mind.

Think Big, Start Small

Begin with a clear view of a risk management approach that fits your business objectives, challenges, and organizational design. Then, carefully target your first projects based on scope (in both business and technology dimensions) and pressing pain to the business.

Small Incidents Can Cause Big Problems

Small incidents such as server failure, cable cut, air conditioning power disruptions, application upgrade, or social engineering (wherein an employee innocently discloses certain company information to a malicious outsider) may cause a big problem for the organization.

Bring Together the IT Silos

One of the major barriers to the first steps in managing IT and infrastructure risk is that there have been so many silos in which IT and infrastructure risk is typically addressed. Risk to a business activity from availability, change, access control, data protection, perimeter security, crisis management, recovery, physical security, project risk, incident management, and more are all managed separately in silos. This makes it very difficult to understand the impact to the business from a range of threats, understand root cause, and prioritize actions.

All-Hazards View of Risk

Ideally, there should be one global enterprise risk profile from the perspective of a specific business activity (e.g., a business unit or single application). Due to complexity and lack of tools and techniques, it is a daunting task to achieve this goal. Yet it is the only way that a business owner can understand the weak links that can cause loss of revenue, fines, penalties, or reputation damage. To work toward this goal, organizations are creating operational risk dashboards to obtain all-hazards views of risk in any given time frame. Often, the first step is just to create a common-view language of risk and common lists of threats and assets (people, process, information, technology, facilities) that could be impacted.

Bring Together Business and IT Leaders for an End-to-End View of Risk

Assuming a more proactive posture on operational risk, the risk management team can begin development of a program to create greater staff awareness of specific

risk drivers and to lay the foundations for risk-informed decision. At the heart of this program is an early warning system (EWS) that monitors performance of key processes and alerts appropriate managers to potential control failures. Early warning indicators (also known as "predictive controls") permit proactive management of operational risk, shifting emphasis away from passive loss recovery to the active prevention of loss events at their source. Early intervention works better (and costs less), with more real-time performance measurement automation.

Make It Simpler

Risk management itself is very simple. Personally, we identify, analyze, plan, and respond hundreds of times a day—even decisions that involve life and death.

What make it complicated are the business environment, organizational design, business process, geographic dispersion (people and facilities), multiple internal IT management models, and range and change in IT hardware and software.

Making it simpler means reducing time and effort for *both* the IT and business people involved. The steps outlined in this chapter can help you immediately start making it simpler to understand and manage risk by applying consistent process across silos in the context of an end-to-end business activity. This not only makes it simpler, but makes it easier to progress risk management projects when a combined risk analysis can be presented to a business line owner with priorities and options for action.

GLOBAL IT STANDARDS MATRIX

With so many IT frameworks in play, it is helpful to matrix them as to their support for compliance, fraud, financials, technology, legal, outsourcing, supply chain, mergers and acquisitions, industrial and electronic espionage, human resources, and the environment.

Exhibit 5.1 provides such a comparative matrix for the more popular standards that impact IT risk management.

COBIT and ITIL are both widely used standards and practices in organizations around the world. Val IT, along with COBIT, was created by the IT Governance Institute and also provides an important contribution to IT risk management at the investment level. With widespread use and great value, this table is designed as a handy-dandy guide to helping you to more easily use them together. Exhibit 5.2 compares these three leading IT risk frameworks.

To conclude this exhibit, an example might help. Consider your car. Val IT helps you prioritize the investments you make in it—new tires, detailing, tune-up, brakes, and such. COBIT provides the general quality controls at a repair shop about a good brake job (but not car make and model specific). ITIL provides general best practices on a tune-up to make them more efficient and consistent. These are great and help you decide how to spend your money, helps the owner of the repair shop give good service, and helps the mechanic work more smoothly. Without these, risk is much higher. The insurers will also be happier. Yet, at the end of the day, your individual make and model car needs a tune-up with specific parts and procedures. For this, you will always need to implement these standards in a way that reflects your IT and

Standard/ Framework	Country of Origin	Description	Compliance	Fraud	Financial	Technology	Legal	Outsourcing	Supply Chain	M & A	Industrial/Economic Espionage	Human Resources	Environment
General Business Risk Framework													
COSO - ERM	USA	Driven by SOX-based risk management using its unique "cube" framework. Detailed appendices for implementation. Created by the Committee of Sponsoring Organizations of the Tread way Commission (COSO) Supported by Institute of Internal Auditors (IIA). CobiT - COSO-ERM mapping document available.	x	x	x	x	x	x	x	x	x	x	x
A Risk Management Standard (ARMS) 2002	UK	Brief and process model driven. No charge public download. Created in the UK by three organizations: Association of Local Authority Risk Manager (ALARM) and Institute of Risk Management (IRM). Supported by Federation of European Risk Management Associations. It has been translated into several languages. Much additional good information at the individual association websites.	x	x	x	x	x	x	x	x	x	x	x
AS/NZS 4360:2004	Australia & New Zealand	Brief and process model driven. Handbook also available provides implementation guidance. Fee charged Used beyond ANZ.	x	x	x	x	x	x	x	x	x	x	x
Open Compliance and Ethics Group (OCEG) Foundation "Red Book"	USA	Is a business level open standard. No charge public download of base document. Provides guidance about the core processes and capability to enhance, culture and address governance risk management and compliance requirements. Beyond the open standard extensive implementation guidance is available.	x	x	x	x	x	x	x	x	x	x	x
ISO/IEC Guide 73, Risk management-Vocabulary-	Multi	Terminology guide for risk management. See also ISO / DIS 31000 Risk management - Principles and guidelines on implementation. Fee charged.											
		Specific-Purpose Risk Framework											
Six Sigma	USA	As a management approach, it was developed at Motorola for quality management, refining earlier techniques developed in several countries. "Littlt" six sigma refers to statistical analysis of variance in a process.	x	x	x	x	x	x	x	x	x	x	x
Control Objective for Information and Related Technology (CobiT®)	Multi	Created by the IT Governance Institute, the Control Objectives for Information and related Tehnology (COBIT) is an internationally-recognized standard for the control of IT processes. Well documented. It's sister standard is ValIT. CobiT has been mapped to many other standards & practices including COSO-ERM. No charge public download. Much additional documentation available.	x	x	x	x	x	x	x	x	x	x	x
Val IT™	Multi	Created by the IT Governance Institute, ValIT provides an approach to measure, monitior and optimise the realisation of business value from investment in IT. It address the risk & return aspects of IT portfolio management. ValIT complements COBIT from a business and financial perspective and will help all those with an interest in value delivery from IT. No charge public download. Much additional documentation available.	x	x	x	x	x	x	x	x	x	x	x
IT Infrastructure Library® ITIL & ISO 20000	UK	From the UK office of Government Commerce, a comprehensive set of practices for IT service management. Developed in conjunction with BS 15000.Later became ISO/IEC 20000. Fee charged.	x	x	x	x	x	x	x	x	x	x	x
ISO 27000 Series	Multi	Begun as British Standard 7799 is a set of standards for IT security management. Later formulated as ISO/IEC 17799-1 and -2. 27001 is the Management System, 27002 is the Code of Practice, 27005 is information security risk management, 27006 is for audit & certification bodies. Fee charged.	x	x	x	x	x	x	x	x	x	x	x
NIST (800-30, 34, 58, 53, 84)	USA	From the US National Institute of Standards & Technology, they address computer security, contigency and other risk management topics. No charge public download.	x	x	x	x	x	x	x	x	x	x	x
DRII/BCI Generally Accepted Practices	Multi	A joint effort of the Disaster Recovery Institute International and the Business Continuity institute. No charge public downlaod.	x	x	x	x	x	x	x	x	x	x	x
BCI Good Practices	UK	Business Continuity Institute recommended practices. Public download at no charge. Also available in German and Italian.	x	x	x	x	x	x	x	x	x	x	x
BS 25999	UK	Based in part on the BCI Good Practices. Formal British Standard Institute release in two parts (designated 25999-1 in 2006 and 25999-2 in 2007). Fee charged.	x		x	x		x	x	x	x	x	x
International Risk Governance Council (IBCC) Risk	Switzerland	Address a range of business and technical risks including biotechnology, carbon capture and nanotechnology. Base document is public and no charge.	x			x	x	x	x	x	x	x	x
BITS Shared Assessment Program	USA	Originated in the financial services industry in the US for supply chain/extended enterprise evaluation on security, BC and DR. Now broaden to other industries & countries. Two parts. Structured information Gathering (SIG) tool and the Agreed Upon Procedures (AUP). Both are public downloads, no charge.	x			x	x	x	x	x	x	x	x

x	Predominantly used	For more information on these standards & practices and the organizations who have created them, please visit www.wiley.com/xxxxxx/barnier_marti_resources
x	Selectively used	
	Little used	

EXHIBIT 5.1 Operational Risk—Global Standards Matrix

	ValIT™	COBIT® 4.1	ITIL® 3
Owner	IT Governance Institute	IT Governance Institute	UK Office of Government Commerce
Purpose	Provides best practices for the end user, providing the means to unambiguously measure, monitor, and optimize the realization of business value from investment in IT.	Provides a comprehensive framework for the management and delivery of high-quality information technology—based services. It sets best pratices for the means of contributing to the process of value creation.	Is an approach to IT service mangement. ITIL is a cohesive best practice framework, drawn from the public and private sectors internationally. It describes the organization of IT resources to deliver business value, and documents processes, functions and roles in.
Audience	IT planning and governance leaders, business line owners, business-IT liaisons, CFO organization, business risk management	IT planning and governance leaders, IT operations, IT-business liaison, IT risk managers, IT auditors	IT service planners, IT service managers, IT operations, IT auditors
Key components	Value Governance, Portfolio Management, Investment Management	Plan & Organize, Acquire & Implement, Deliver & Support, Monitor & Evaluate	Service Strategy, Service Design, Service Transition, Service Operation, Continual Service Improvement
In short	How to invest in IT	How to manage IT operations	How to operate IT
How it helps you	If you are trying to understand risk & return in your set of IT investments, then this is the leading approach for you.	If you are seeking a control framework to measure IT outcomes as a means to improve return, reduce risk or improve quality, then this is the leading approach for you.	If you are seeking consistency, efficiency, reduced process errors and improved quality of IT service delivery, then this is the leading approach for you.
Related disciplines	Should be closely linked with financial management best practices for capital budgeting and portfolio management. Also with project management methods.	Should link upward to ValIT and downward to ITIL. Horizontally, can link with business risk management and control techniques (e.g., COSO).	Should link upward to COBIT. For more detailed information users can turn to domain specific standards for disaster recovery, information security and such. The most detailed guidance comes from best practices for using systems management software and configuring specific IT resources.

EXHIBIT 5.2 Comparison among Val IT, COBIT, and ITIL

business environment, your business objectives, your culture, and your leadership style. Taking this approach helps create a risk-aware governance pattern that works for you, not chokes you.

LINKS TO IT RISK ASSOCIATIONS AND AGENCIES

This chapter is designed to provide an introduction to IT risk management. For further information, there is a wealth of publicly available information via Internet links to the many associations and government agencies that are actors in IT risk management.

ORGANIZATIONS

Committee of Sponsoring Organizations (COSO)
www.coso.org/
Association of Insurance and Risk Managers
www.airmic.com/
Association of Local Authority Risk Managers
www.alarm-uk.org/
Institute of Risk Management
www.theirm.org/
Standards Australia
www.standards.org.au/
Standards New Zealand
www.standards.co.nz/
Open Compliance & Ethics Group (OCEG)
www.oceg.org/
International Standards Organization (ISO)
www.iso.org/iso/home.htm
International Society of Six Sigma Professionals
www.isssp.com/
IT Governance Institute
www.itgi.org
Information Systems Audit & Control Association
www.isaca.org
U.K. Office of Government Commerce (OGC)
www.ogc.gov.uk/guidance_itil.asp
U.S. National Institute of Standards and Technology Computer Security Resource Center
www.ogc.gov.uk/guidance_itil.asp
Disaster Recovery Institute International
www.drii.org
Business Continuity Institute
www.thebci.org
British Standards Institute
www.bsi-global.com/en/Standards-and-Publications/
International Risk Governance Council
www.irgc.org/
BITS Financial Services Roundtable
www.bitsinfo.org

SELECTED STANDARDS & PRACTICES

There are a number of other helpful standards and practices, including those created on a country or industry basis. This listing provides a selection of guidance.

COSO ERM Integrated Framework
www.coso.org/ERM-IntegratedFramework.htm

A Risk Management Standard
From AIRMIC web site
www.airmic.com/en/Library/Risk_Management_Standards/
From ALARM web site
www.alarm-uk.org/PDF/rmstandard.pdf
From IRM web site
www.theirm.org/publications/PUstandard.html

ORGANIZATIONS

AS/NZS 4360:2004 Set (includes handbook)
www.saiglobal.com/shop/Script/Details.asp?DocN=AS564557616854
OCEG Governance Risk and Compliance Foundation
www.oceg.org/View/Foundation
ISO Publication 73
www.iso.org/iso/catalogue_detail?csnumber=34998
Motorola University for Six Sigma
www.motorola.com/motorolauniversity.jsp
Val IT
www.isaca.org/Template.cfm?Section=Val_IT4&Template=/ContentManagement/ContentDisplay.cfm&ContentID=39994
COBIT
www.isaca.org/Template.cfm?Section=COBIT6&Template=/TaggedPage/TaggedPageDisplay.cfm&TPLID=55&ContentID=7981
ITIL
www.itil-officialsite.com/Publications/Core.asp
ISO 27000 series
www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=42103
NIST (especially 800-30, 34, 58, 53, 84)
http://csrc.nist.gov/publications/PubsTC.html
DRII/BCI Generally Accepted Practices
www.drj.com/GAP/
BCI Good Practices
www.thebci.org/gpgdownloadpage.htm
BS 25999-1
www.bsi-global.com/en/Shop/Publication-Detail/?pid=000000000030157563
BS 25999-2
www.bsi-global.com/en/Shop/Publication-Detail/?pid=000000000030169700
IRGC Framework Introduction
www.irgc.org/IMG/pdf/An_introduction_to_the_IRGC_Risk_Governance_Framework.pdf
BITS Shared Assessment Program SIG and AUP
www.bitsinfo.org/FISAP/index.php

DOWNLOADABLE REFERENCE

IBM® Tivoli® Unified Process (ITUP) is a Web-based tool that provides detailed documentation based on industry best practices such as ITIL and COBIT.
www-306.ibm.com/software/tivoli/governance/servicemanagement/itup/tool.html

CHAPTER 6

An Operational Risk Management Framework for All Organizations

Anthony Tarantino, Ph.D.

INTRODUCTION

Risk is usually defined as the possibility of a loss or injury created by an activity or by a person. Risk management frameworks attempt to identify, assess, measure risk, and then develop countermeasures to mitigate its impact. This typically does not aim to eliminate risk for there is little opportunity without some degree of risk. An organization that is too risk adverse will not be likely to attract investors.

The types of risks that impact organizations vary depending on such factors as the region, industry, and level of globalization. Banks worry about credit and market risks. Insurance companies worry about actuarial risk. Many firms worry about reputation, and legal risks. Risks can be internally or externally based, but one area of risk impacts all organizations—operational risk. This is true for public and private companies, nonprofits, and government agencies.

The growing losses from the current financial liquidity crisis which brought about the meltdown of the subprime mortgage market demonstrate a catastrophic failure in operational risk management. While some are still arguing that it is also a failure of credit risk, this misses the root cause. Credit risk is not a primary cause when a lender intentionally loans money to borrowers unable to qualify for traditional loans, when all normal due diligence in checking credit and employment histories are ignored or even intentionally falsified, and when the lender conspires with appraisers to inflate property values.

In the past operational risk has not been a major area of concern for most financial service institutions. Outside of financial services, it has received even less attention with most firms focused on market risks and opportunities. This is ironic given that operational risk failures are behind most of the marquee scandals of the last two to three decades. Also ironic is that the recent scandals and crises in subprime and Société Générale occurred in institutions with some of the most sophisticated and robust operational risk management protocols.

DEFINITION AND CATEGORIZATION OF OPERATIONAL RISK

The banking and insurance industries are addressing operational risk in a major way with new capital adequacy accords known, respectively, as Basel II and Solvency II. This is no academic exercise, requiring institutions to reserve capital to cover their operational risks. The Basel Committee of the Bank for International Settlements (BIS) and the Solvency Committee of the International Association of Insurance Supervisors (IAIS) define operational risk as the risk of losses resulting from inadequate or failed internal processes, people and systems or from external events. Although designed for financial institutions, this definition should be applicable for any industry, institution, or individual.

The Basel and Solvency approach to operational risk breaks it into seven major categories, 18 secondary categories, and 64 subcategories. The great majority is not unique to financial services and can provide a good framework for addressing operational risk in any industry:

1. **Internal Fraud**
   - **a.** Unauthorized Activities
     - **1)** Transactions not reported (informational)
     - **2)** Transaction type unauthorized (with monetary loss)
     - **3)** Mismarking of position (international)
   - **b.** Theft and Fraud
     - **1)** Fraud/credit fraud/worthless deposits
     - **2)** Theft/extortion/embezzlement/robbery
     - **3)** Misappropriation of assets
     - **4)** Forgery
     - **5)** Check kiting
     - **6)** Smuggling
     - **7)** Account takeover/impersonation/etc.
     - **8)** Tax noncompliance/evasion (willful)
     - **9)** Bribes/kickbacks
     - **10)** Insider trading
2. **External Fraud**
   - **a.** Theft and Fraud
     - **1)** Theft/robbery
     - **2)** Forgery
     - **3)** Check kiting
   - **b.** System Security
     - **1)** Hacking damage
     - **2)** Theft of information (with monetary loss)
3. **Employment Practices**
   - **a.** Employee Relations
     - **1)** Compensation, benefit, termination issues
     - **2)** Organized labor activities
   - **b.** Safe Environment
     - **1)** General facility (e.g., slip and fall)
     - **2)** Employee health and safety rules, events
     - **3)** Workers' compensation

- **c.** Diversity and Discrimination
  - **1)** All discrimination types (racial, sexual, sexual orientation, religions, etc.)

4. Clients, Products, and Business Processes

- **a.** Suitability, Disclosure, and Fiduciary
  - **1)** Fiduciary breaches/guideline violations
  - **2)** Suitability/disclosure issues (Know Your Customer, etc.)
  - **3)** Retail consumer disclosure violations
  - **4)** Breach of privacy
  - **5)** Aggressive sales
  - **6)** Account churning (excessive buying and selling of securities by a broker to generate commissions)
  - **7)** Misuse of confidential information
  - **8)** Lender Liability
- **b.** Improper Business or Market Practices
  - **1)** Antitrust
  - **2)** Improper trade/market practices
  - **3)** Market manipulation
  - **4)** Insider trading (on firm's account)
  - **5)** Unlicensed activity
- **c.** Product Flaws
  - **1)** Product defects (unauthorized, etc.)
  - **2)** Model errors (poor design)
- **d.** Selection, Sponsorship and Exposure
  - **1)** Failure to investigate client per guidelines
  - **2)** Exceeding client exposure limits
- **e.** Advisory Activities
  - **1)** Disputes over performance of advisory activities

5. Damage to Physical Assets

- **a.** Disaster and Other Events
  - **1)** Natural disaster losses
  - **2)** Human losses from external sources (terrorism, vandalism)

6. Business Disruptions and System Failures

- **a.** Systems
  - **1)** Hardware
  - **2)** Software and middleware
  - **3)** Telecommunications
  - **4)** Utility outage/disruptions (failures in business continuity)

7. Execution Delivery and Process Management

- **a.** Transaction Capture, Execution, and Maintenance
  - **1)** Miscommunication
  - **2)** Data entry, maintenance, or loading error
  - **3)** Missed deadline or responsibility
  - **4)** Model/system misoperation
  - **5)** Accounting error/entity attribution error
- **b.** Monitoring and Reporting
  - **1)** Failed mandatory reporting obligation
  - **2)** Inaccurate external report (loss incurred)

c. Customer Instate and Documentation
  1) Client permissions/disclaimers missing
  2) Legal documents missing/incomplete
d. Customer/Client Account Management
  1) Unapproved access given to accounts
  2) Incorrect client record (loss incurred)
  3) Negligent loss or damage of client assets
e. Trade Counterparties
  1) Nonclient counterparty performance
  2) Miscellaneous nonclient counterparty disputes
f. Vendors and Suppliers
  1) Outsourcing
  2) Vendor disputes

HOW AUDITORS AND REGULATORS APPROACH RISK MANAGEMENT

No matter what risk framework an organization deploys, it will have to satisfy auditors and regulators, who will typically use the following framework:

- Identity business processes, especially those impacting financial reporting.
- Identity the risks associated with each process.
- Identify the internal controls used to mitigate the risks for each process.
- Create a hierarchy of business processes, risks, and controls.
- Identify the tests to be used in determining the effectiveness of the internal controls.
- Test the internal controls and publish findings.
- Provide an opinion as to the effectiveness of the controls.
- If the controls are found to be ineffective, recommend changes (remediations) and retest the controls.
- Create and maintain a documentation library of the processes, risks, controls, tests, findings, remediations, etc. involved in the risk/control process. This would include a risk/control matrix, process narratives, process flow charts, test procedures, and so on.
- If the internal controls are found to be effective, business owners, internal auditors, and external auditors will sign off on them as part of a certification process. With a few notable exceptions such as France and Canada, most national regulations and audit protocols are based on a COSO framework. The Committee of Sponsoring Organizations of the Treadway Commission (COSO) framework has been in wide usage for many years, but has not lived up to expectations in improving risk management over internal controls. This is due to its lacking even a basic means to quantify risk along these lines. The terrible failures in operational risk management that sparked the financial liquidity crisis of 2007 and 2008 are a stark reminder that all the expensive internal control reforms under the Sarbanes-Oxley Act's (SOX) Section 404, did little to prevent or warn of the problem. SOX section 404 references COSO as an acceptable framework, and most organizations and audit firms have embraced it.

HOW RATING AGENCIES EVALUATE OPERATIONAL RISK

After satisfying regulatory and audit masters, organizations will want to consider how the major rating agencies evaluate operational risk. While satisfying regulators and auditors can keep an organization out of hot water, satisfying rating agencies can offer real monetary benefits in lower capital and insurance costs. The rating agencies such as Moody's, Fitch, and Standard & Poors have published various white papers and standards as to what they will look for in a well risk-managed organization. They include the following elements that are applicable across industries:

- A risk management committee and working groups with an enterprise-wide charter, which possesses the needed training, expertise, resources, and time to do its job. (They do not translate this to include a risk committee reporting to the board of directors, which we advocate in the next section.)
- A risk identification process that is enterprise-wide, independently reviewed, and audited on a periodic basis. (The frequency typically mandated by external auditors may not be adequate for areas of high risk.)
- Assurances that the risk committee and risk managers communicate on a regular basis beyond the reporting of risks. (This would be greatly facilitated by a risk committee reporting to the board of directors.)
- A risk-weighted approval process for new products and strategies. (We provide an example of how this can work in the next section.)
- An ongoing effort to diversify risk on an enterprise-wide basis. (The goal is to prevent an overconcentration of risk in any one area that could jeopardize the health or very existence of the organization.)
- A centralized and dedicated risk management organization that is staffed with the appropriate subject matter experts and has the charter to remain independent from those taking the risks. This organization would be charted to identify, communicate, and audit all risks without fear of retaliation. (This process clearly failed in most of the leading financial services organizations as risk managers were ignored or punished for raising concerns over the subprime market.)

AN OPERATIONAL RISK FRAMEWORK FOR ALL ORGANIZATIONS

We offer an approach to operational risk management that can work for both large and smaller organizations. It requires no sophisticated analytical tools or large technology investments. It starts by ranking each of the 64 subcategories of operational risk described previously by three criteria:

1. Its financial impact
2. The ability to detect it
3. Its likelihood of occurring

Applying a simple one-to-five rating for each of these criteria to each of the 64 subcategories and then adding them together can provide a convenient means to

prioritize operational risk management efforts. The Italian economist Pareto developed an 80–20 rule that works with few exceptions. In this case, about 10 to 15 of the 64 categories will probably represent at least 80 percent of an organization's risk exposure. These would be the risks that an organization should focus on in terms of creating countermeasures in business process and technology improvements.

Using Six Sigma black belts that are highly trained in root cause analysis, problem solving, and listening to the voice of the customer will make this process much more effective. Existing resources can be cross-trained in Six Sigma and require little administrative overhead or technology investments. It would be prudent to include both business and technology resources as part of the Six Sigma training program. This will provide a better balance in assessing the relative importance of process improvements and technology improvements.

The Basel and Solvency committees and rating agencies have acknowledged Six Sigma as a best practice framework in operational risk management. Providing these Six Sigma black belts with a Lean perspective would be even better. Lean is the popular name for a philosophy that strives to eliminate waste of all types. It was developed by the Toyota Corporation and came to the United States and European Union as Just-In-Time (JIT) manufacturing; it has evolved to beyond manufacturing.

For the 20 percent of high-risk areas that represent the great majority of an organization's total risk, it would be prudent to use Six Sigma black belts to develop the means to automate the controls over these risks. The higher the level of automation the better, unless it can be shown not to be cost effective in mitigating risks. Typically, manual controls are not desirable, while automated controls provide higher levels of protection. Among automated controls, preventative controls are more desirable than detective controls. The highest level of automated preventative controls should include a system of hierarchical dashboard notifications and alerts when controls are breached or threatened. Regulators, rating agencies, and auditors will typically reward organizations with automated preventative controls. They recognize that manual controls are typically ineffective and require more frequent and costly auditing than automated controls.

For organizations with the capabilities of capturing their history of operational losses, this data can be used to help weigh the 64 subcategories. With this methodology, data are transformed into loss frequency and severity distributions. The major issue in using historical loss data is that thousands of loss events are required to develop modeling. Outside of financial services, it is not typical for organizations to spend the considerable resources and technology investments to normalize, categorize, evaluate, and model such loss data. Modeling of operational risk is typically reserved for the financial services industry as a means to calculate economic or regulatory capital.

It may be advisable to create a more granular categorization than the 64 offered here for the handful of risks that are of greatest concern. For example, external theft may be too generic a category to provide the right focus. An organization may face a variety of theft threat types.

Beyond utilizing Six Sigma black belts to attack risk exposure, organizations should consider increasing influence of risk management at the board of directors. Most country laws now require an audit committee made up of independent directors and financial experts to report to the board of directors. With the exception of financial services, these audit committees also are looked upon for risk management

oversight. The skill sets of financial experts and risk managers overlap somewhat, but are not the same. A risk committee made up of risk experts and with a majority of independent directors reporting directly to the board would be a good way to better balance opportunities and risks. Ideally, they would closely coordinate their activities with the audit committee. This process will work best in an organization in which the positions of chief executive officer (CEO) and chairman of the board (CoB) are held by separate individuals. The subprime crisis clearly demonstrated how risk experts were ignored by an all-powerful company head holding both positions. With the unrelenting pressure to satisfy investors, it is not realistic to expect one individual, no matter how talented, to balance risks and opportunities.

CONCLUSION

In summary, we believe that a basic operational risk framework can be deployed by all organizations. Organizations can scale this framework according to their capabilities and requirements, including more sophisticated risk management tools. Our framework for operational risk would include:

- Use the 64 subcategories developed by the Basel and Solvency committees to rank operational risk.
- Apply a simple one-to-five ranking for each as to its likelihood, delectability, and financial impact.
- Focus efforts on those areas with the highest risk scores.
- Cross-train business and IT resources in Six Sigma.
- Apply Six Sigma problem-solving resources and techniques to improve risk mitigation.
- Increase automation of controls over the areas with the highest risk.
- Create a risk committee with a majority of independent directors and risk experts reporting to the board of directors.

CHAPTER 7

Financial Risk Management in Asia

Anthony Tarantino, Ph.D.

INTRODUCTION

Asia has become a major force in global trade creating many critical interdependencies with Western economies. As such, Asia presents significant opportunities and risks for its trading partners and for investors. This chapter discusses some of the major areas of risk for the leading Asian economies and offers the means to mitigate the risk.

As measured by purchasing power parity (PPP), China's economy will be larger than the United States' by the middle of this century, and India's will be roughly the same size as the U.S. economy.[1] China is also becoming a major financial powerhouse, owning over 19 percent of U.S. Treasury securities, with $518 billion as of July 2008. (Japan is the largest owner, with 22 percent or $593 billion).[2] China is now the second-largest trading nation behind Germany and is running a trade surplus of roughly $30 billion. Its major trading partners include Japan, the United States, South Korea, and Germany.[3]

While India is not nearly as large as China in global trade (about 1.2 percent of world trade as of 2006 per the World Trade Organization), it has become a key player in providing critical information technology (IT) outsourcing and infrastructure to support a wide variety of organizations in the EU and the United States.[4] International Business Machines Corporation's (IBM's) explosive growth in India is a good example of its critical role in IT services and infrastructure. From 2003 and 2007, IBM's Indian employee head count grew by over 800 percent—from 9,000 to 74,000. IBM is now the largest multinational in India.[5] IBM's growing reliance on India is important because IBM is by far the largest IT provider to the banking and financial services industries and as such handles massive amounts of financial transactions throughout the world.

So China and India are now critical to the global economies as exporters of materials and IT services. China is also critical to world financial markets because of its heavy stake in U.S. Treasuries.

Another area of risk comes for investors in Asia's securities markets. These markets have become very attractive to investors over the last decade, but have suffered major corrections as the speculative bubble burst in late 2008. India, China, and Japan have programs under way to implement Sarbanes-Oxley (SOX)-like regulations to improve transparency and accountability in financial reporting. We will

argue that these reforms will not provide adequate risk transparency and that the nature of their cultures will frustrate whistle-blowing, auditing, and business press coverage. These are some of the reasons for concern:

- *SOX-like regulations do not translate into risk transparency.* The Sarbanes-Oxley Act and other financial transparency initiatives do little to provide risk transparency. There is no viable means in their Committee of Sponsoring Organizations of the Treadway Commission (COSO) framework to rationalize, score, and rank risks. (This is discussed in detail in Chapter 24 on corporate governance.) Unlike Western societies, there are few whistle-blower protections and little history of whistle-blower activities. As we detail in our *Manager's Guide to Compliance,* whistle-blowers are the primary vehicle for uncovering corporate wrongdoing. Internal and external audits are not nearly as effective.
- *Auditors lack deep expertise.* Auditors in Asian countries may lack the experience and do not perform at the level of detail as their western counterparts. Japanese auditors follow the generally accepted accounting principles (GAAP) based on U.S. GAAP, but typically expend half the man-hours in performing financial audits. The levels of expertise and level of detail effort are much lower in other Asian countries.
- *The International Financial Reporting Standards (IFRS) transition introduces risks.* Like much of the world, Asian countries are adopting the IFRS, which is a much-needed reform and standardization. As we discuss in Chapter 9, this transition will introduce risk. The principle-based IFRS does not come with an easy to follow checklist of rules found in the U.S. and Japanese GAAP. Less experienced accounting professionals will struggle to agree on the guidelines. Fraudsters will see this as an opportunity to cheat.
- *A vigorous investigative news media is lacking.* A free and competent business news media is critical in identifying corporate wrongdoing and weaknesses. Many Asian countries lack Western levels of business press acumen and freedom. The importance of this cannot be overemphasized, as most every marquee scandal has been exposed by the press and not by government regulators, auditors, and rating agencies. Watergate and Enron are classic examples of the critical role journalists play as a fourth branch of government, providing checks and balances among executive, legislative, and judicial branches of government.

The following is a summary status of the news media in major Asian countries from Freedom House[6]:

- *Malaysia.* News media are constrained by significant legal restrictions and intimidation, which was intensified in 2006 by government attempts to suppress public discussion of divisive and potentially explosive issues. Malaysian law requires all the print media to obtain yearly permits, which can be revoked by the government without any judicial review.
- *Indonesia.* Legal intimidation against journalists restricts investigative reporting. Laws carrying criminal penalties prohibit insulting public authorities and state institutions, and the direct relay of foreign broadcast content by local private radio and television stations. Other major problems are the continued violence against journalists, and the government's continued ban on foreign journalists entering West Papua.

- *India*. The news media continue to be robust and by far the freest in South Asia, but journalists still face many constraints. India passed a Right to Information Law in 2005. An independent journalist body, the Press Council of India, acts as a self-regulatory mechanism investigating complaints of journalistic misconduct. Violence against investigative journalists continues to be a problem, including the murder of one journalist who exposed an official's misconduct and corruption. The threat of violence has led to self-censorship. Much of the print media are privately owned and provide diverse coverage, frequently scrutinizing the government.
- *China*. The trends are negative with increased crackdowns on journalistic freedoms. While the Constitution guarantees freedom of speech, assembly, association, and publication, other regulations restrict these rights to the national interest, as defined by government-controlled courts. The Communist Party maintains direct control over the news media around topics deemed to be politically sensitive. New regulations were introduced in 2006, controlling the distribution of foreign media coverage of unforeseen events. The crackdown was sparked by a growing chorus of public protests, the growth in online news availability, and the need for the press to become profitable. While the Chinese media are state owned, most no longer receive state subsidies and must rely on advertising revenue. This has shifted their loyalties away from the government and toward their readers.

While this chapter focuses on Asia, much of the regulatory and risk environment we describe applies to Latin America (covered in Chapter 8), Islamic nations, and Africa. The following areas of concern apply to all these regions: the lack of strong accounting principles, weak regulatory oversight, closely held corporate ownership, ineffective and repressed news media, lackluster corporate boards, and a culture where fraud and corruption are commonplace.

RISKS IN ASIAN SUPPLY CHAINS

Asian exports continue to expand, and its 10 largest exporters now comprise 35 percent of global exports. As shown in Exhibit 7.1, China ranks second behind Germany and ahead of the United States.[7]

Because of Asia's major role as an exporter, it presents significant supply chain risks to its customers. Matt Eikington, who manages the risk practice for Marsh Consulting, surveyed his clients in 2006, who listed the following risk areas as a high priority or major focus[8]:

- Infrastructure Risks: 44%
- Quality and Counterfeiting: 44%
- Ethical Risks: 40%
- Regulatory Risks: 40%
- Financial Risk: 40%
- Fraud and Corruption: 30%
- Natural Disasters: 28%

EXHIBIT 7.1 2007 Asian Exports

Global Country Rank	Country	Exports (2007 est.)	Asia's Cum. Total of Global Exports
N/A	World	$13,890,000,000,000	N/A
2	China	$1,220,000,000,000	8.8%
4	Japan	$678,100,000,000	13.7%
10	South Korea	$379,000,000,000	16.4%
12	Hong Kong	$345,900,000,000	18.9%
14	Singapore	$302,700,000,000	21.1%
17	Taiwan	$246,500,000,000	22.8%
21	Malaysia	$176,400,000,000	27.7%
25	India	$151,300,000,000	32.6%
26	Thailand	$151,100,000,000	33.7%
31	Indonesia	$118,000,000,000	34.5%

Given the major price advantages from Asian exporters, there is no simple means to mitigate these risks. The classical procurement approach is to develop alternative suppliers who provide more stability than exporters from developing economies. Unfortunately, there are no viable alternative suppliers in many cases. The United States and European Union (EU) have seen a major erosion of their manufacturing bases over the past two decades. Even government efforts to subsidize their domestic industries have failed to stem the tide.

The first recommendation is to become very familiar with critical suppliers from Asia and other nontraditional areas. This is much more than conducting occasional site visits. It entails performing a comprehensive due diligence around the supplier's financial stability, reputation, IT security and infrastructure, quality control, and disaster recovery (business continuity) plans.

The recent Chinese scandals (toys contaminated with lead paint, and melamine poisoning in several food products) are stark evidence of the danger of taking a hands-off approach. In both cases, the Chinese government regulations were in place, but a culture of greed and corruption overwhelmed regulatory oversight.

A second recommendation is to consider taking a less aggressive approach to inventories than Just-in-Time (JIT) would suggest. JIT is a major improvement over traditional approaches as perfected by Taiichi Ohno for Toyota, but JIT is designed for very reliable suppliers who are physically, culturally, and politically close to their suppliers. There is a trade-off between higher carrying costs and lower prices that Asian exports provide. Japan's keiretsu, the foundation of Toyota's JIT manufacturing system, created a very reliable supplier base and is based on very tightly interlocking business relationships and shareholdings. It is risky to assume the same level of service from distant suppliers without a proven partnership relationship.

A third recommendation in evaluating suppliers from various regions is to use the World Bank's database of corporate governance. While this will not guarantee the performance of any one supplier, it will provide valuable insights into the overall governance culture for a given country. Exhibits 7.2, 7.3, 7.4, and 7.5 show four of the World Bank's six categories of governance for the leading Asian economies.

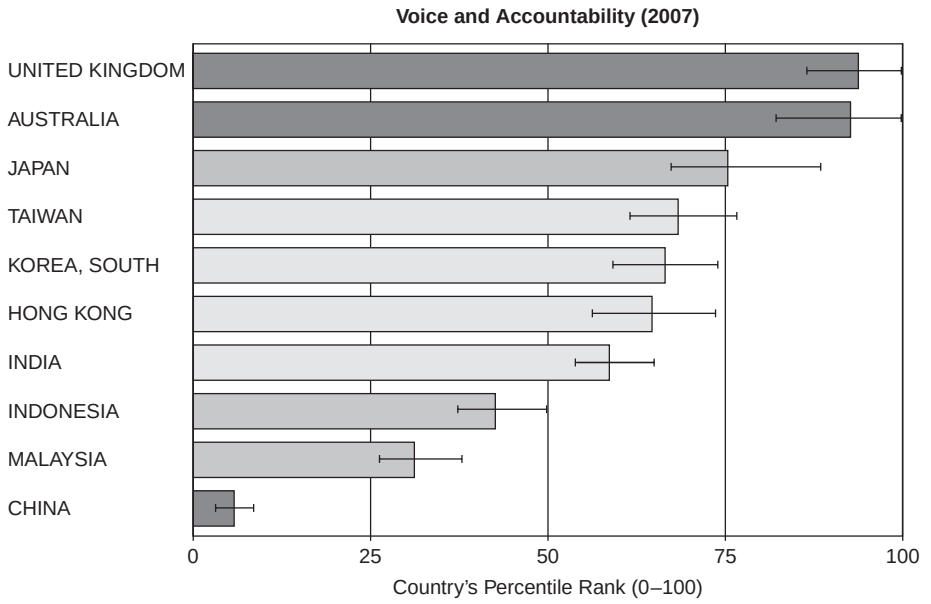


EXHIBIT 7.2 World Bank, Voice and Accountability for Asian Countries

They are compared to the United Kingdom, which has consistently ranked among the world's best-governed economies.

The very low scores China receives in voice and accountability indicates that there is no culture that would encourage whistle-blowers to come forward even during situations that endangered the health and lives of their own children—the melamine poisoning has left 13,000 infants hospitalized and four dead.[9] In the United

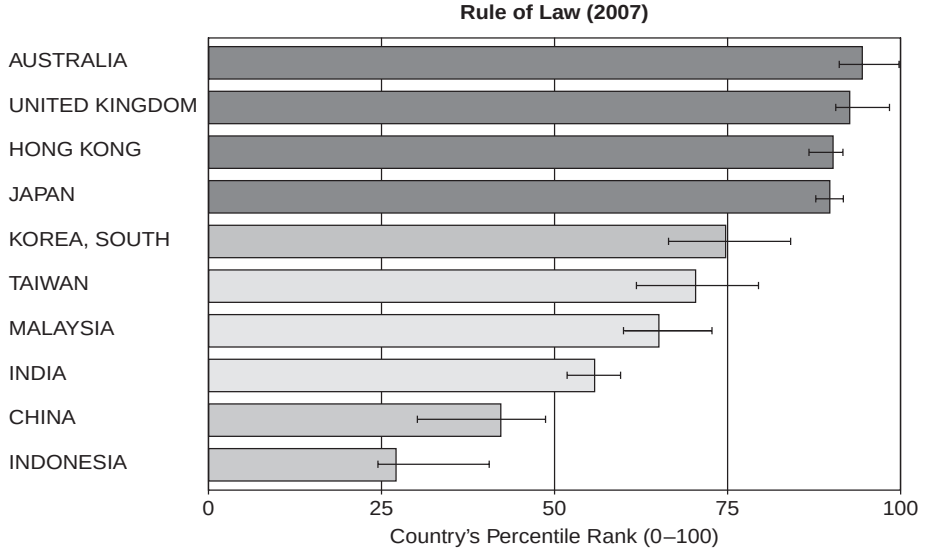


EXHIBIT 7.3 World Bank, Rule of Law for Asian Countries

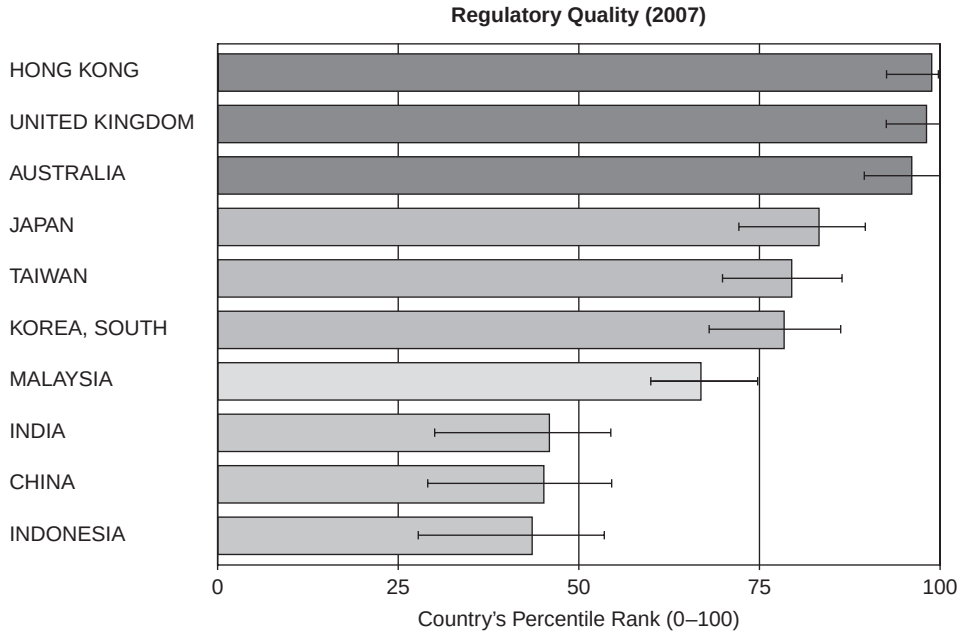


EXHIBIT 7.4 World Bank, Regulatory Quality for Asian Countries

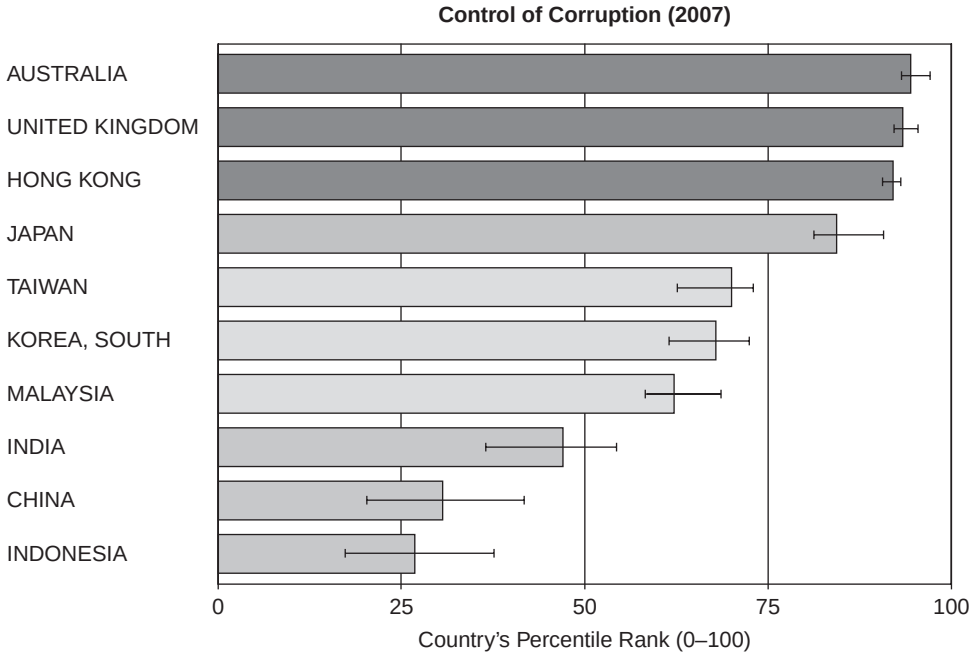


EXHIBIT 7.5 World Bank, Control of Corruption for Asian Countries

States, whistle-blower protections are posted in employee workplaces. Whistle-blowers have the confidence that they can report abuses and wrongdoing to government officials and/or go to the news media. In China, government officials are often part of the corruption, and the news media are controlled by the government. So even the most sincere and courageous whistle-blower intentions are unlikely to bear fruit.

India, Indonesia, and China are consistent in their low governance scores compared to such East Asian countries as Taiwan, South Korea, and Malaysia.

Besides the ethical and moral issues at play, such supplier activities expose their customers to very high financial risks. U.S. and EU tort laws invite major lawsuits in such situations that will take years to play out in courts. Juries and judges are not likely to accept pleas of ignorance of such abuses.

RISKS IN ASIAN FINANCIAL MARKETS

Asian financial markets have become very attractive to investors because of very high growth rates as compared to Western financial markets. In our *Governance, Risk, and Compliance Handbook*, we compare corporate governance levels and gross domestic product (GDP) growth rates. We conclude that higher governance ratings do not translate into high GDP growth rates. This has contributed to very attractive growth rates, at least until the global financial crisis hit. The risks in these markets is now obvious, as they declined at much higher rates than their U.S. or European counterparts.

Exhibit 7.6 uses Yahoo! Finance online charts to compare the growth rates of the India's Bombay Index (SENSEX), Shanghai Composite Index (CH,SHI0), Hong Kong's Hang Seng Index (HK,HIS), Japan's Nikkei 225 Index (JP,N225), and Great Britain's Financial Times Stock Exchange (GB,FTSE). While Asian investors enjoyed

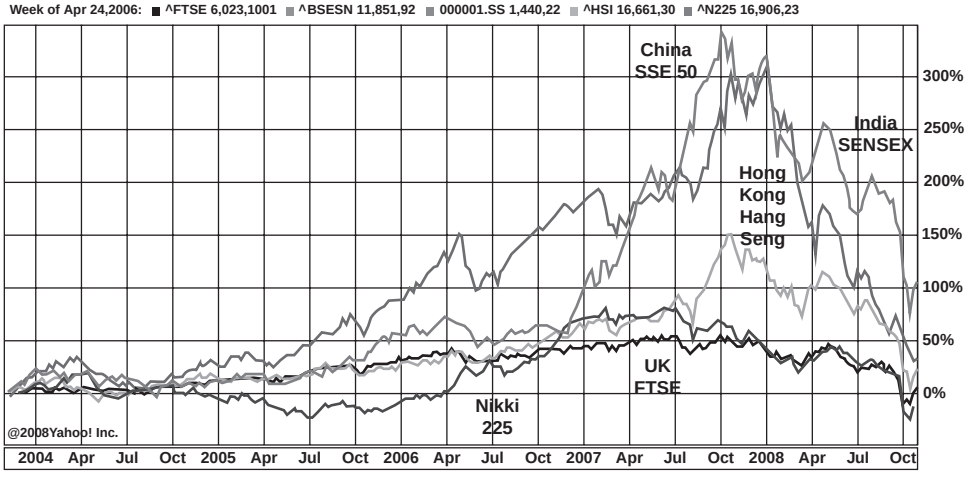


EXHIBIT 7.6 Yahoo! Finance, Five-Year Performance of Major Asian Stock Indexes vs. London Stock Index

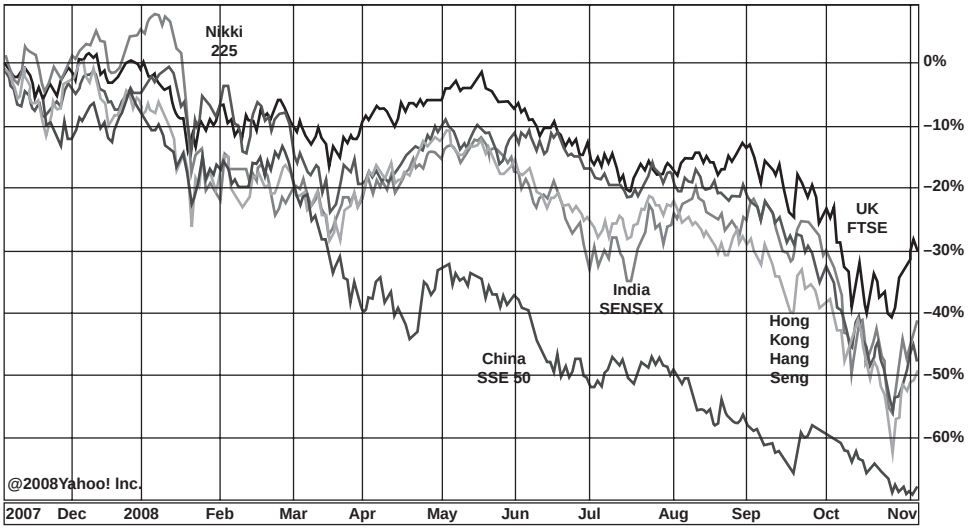


EXHIBIT 7.7 Yahoo! Finance, One-Year Performance of Major Asian Stock Indexes vs. London Stock Index

extraordinary growth from 2006 to late 2007, they suffered huge losses in 2008. Britain's FTSE, in contrast, has shown much more stability during periods of market turmoil. For long-term investors, over five years in our comparison, investors in India's SENSEX would have enjoyed about a 100 percent gain, those in China only about a 30 percent gain, those in Hong Kong only about a 20 percent gain, and those in the United Kingdom would have made no gain.

Exhibit 7.7 uses Yahoo! Finance online charts to show that investors in Asian markets who entered the market 12 months ago would have suffered significantly higher losses than in U.S. or U.K. markets. The Shanghai index lost about 70 percent, while Bombay, Hong Kong, and Japan lost between 40 percent and 50 percent. Investors in the United Kingdom fared the best with a loss of about 30 percent.

We compare Asian markets to the United Kingdom, because of the stark differences in their levels of corporate governance. The United Kingdom has consistently enjoyed the highest corporate governance scores according to the World Bank, while China has scored very poorly. The lack of a strong governance framework means that markets lack financial transparency and accountability. Therefore, investors have poor visibility into the actual performance of the companies they invest in as compared to their Western counterparts.

The following are the highlights of efforts to improve financial transparency and accountability in some of the leading Asian economies. We also include an assessment of the risks that will remain.

China[10]

China is beginning the process of adopting SOX-like regulations aimed at improving financial transparency and accountability for companies and their auditors. An annual survey by Deloitte conducted in 2008 found that less than half of companies

had created a system of internal controls, but that over half of all companies surveyed had either no internal controls or inadequate controls.[11]

China's Ministry of Finance, China Securities Regulatory Commission, the National Audit Office, China Banking Regulatory Commission, and China Insurance Regulatory Commission joined together in June 2008 to announce a new Basic Standard for Enterprise Internal Control. Chinese-listed companies will have until July 2009 to comply. This is a very short time frame considering the magnitude of the change that is required for companies, their regulators, and their auditors.

Chinese companies will face the following obstacles in implementing internal controls under China's SOX framework:

- Controls are not standardized or automated across the enterprise.
- Company executives are not fully committed to the process.
- IT infrastructure and IT personnel are inadequate to provide robust data governance, storage, and access required for internal controls.
- There a few internal or external resources available that fully comprehend the creation, documentation, and audit of internal controls.
- Internal controls are treated as a necessary evil to pass audits and not part of company operations.

While this is good news in improving corporate governance, investors and other stakeholder need to proceed cautiously. Sarbanes-Oxley was a very painful and expensive process for U.S.-listed companies, but the improved internal controls, financial transparency, and accountability did little to warn of the meltdown in the financial services industry. While China has made tremendous progress in improving its financial acumen, its accounting, auditing, and regulating skills are not yet up to Western standards.

Another note of caution: Financial transparency is counter to traditional Asian approaches to business in which such transparency is seen as giving away strategic information to one's competitors. China has a tradition of imposing regulations but not providing the mechanisms and political will to enforce them. This has been a major factor in the recent food and toy poisoning scandals.

Japan[12]

Effective in May 2006, Japan has imposed a new corporate law. It holds company boards responsible for development and implementation of internal controls—similar to U.S. SOX section 404. It also requires company management to introduce and utilize internal control to meet regulatory requirements. The levels of internal controls are not detailed in the company law and will require a consensus between companies, audit firms, and regulators. This is similar to the International Standard of Audit 315, Planning and Preparation for an Audit. It also requires an internal controls audit standard similar to the U.S. Public Company Accounting Oversight Board's (PCAOB) Audit Standard Number Five.

Japan also passed a Financial Instruments and Exchange Law that went into effect in July 2006. It tightens regulations on fraudulent financial reports, misinformation, and manipulation of stock price criminal penalties (10 years of prison and/or 10 million yen fines). This is similar to U.S. Title VIII and IX (Sections 802,

807, 903, 906, and 1105). Effective April 2008, a company's CEO, COO, and CFO are required to certify the accuracy and completeness of financial reports. Finally, it requires a company's auditor to independently assess financial reports.

These regulations are now commonly referred to as JSOX. Japan faces the following issues in implementing them:

- *CEO-appointed Auditors System.* This is a conflict of interest in that most of the auditors are not independent. Auditors had been empowered by changes in the former Commercial Law, but it was still difficult to say that the Auditors System functioned well.
- *Problems with the CPA Auditing System.* The Kanebo case and other ethical failures forced changes to Exchange and CPA laws with new CPA penalties. One major audit firm suspended for two months and a midsized firm was disbanded.
- *Audit efforts.* Audit working time in Japan is about one third to one half of the audit working time in America, England, and Germany. CPA efforts will have to increase greatly to meet JSOX requirements.
- *Audit firm rotation.* The JICPA has also decided to introduce the rotation system, so that CPAs now are not allowed to take charge of the same client for more than five consecutive fiscal years.

India[13]

Corporate governance in India has been hindered by the ethical atmosphere of publicly traded companies in which there been little accountability and, until recently, there have been no restrictions on the level of independence of the board of directors. The main problem comes from the lack of independence of the board of directors. Directors who are not independent can make choices to benefit themselves rather than the company and the shareholders.

India is still in its early stages of development, but hopefully with models like the U.S. Sarbanes-Oxley Act and the United Kingdom's Cadbury Code, and now India's Clause 49, they will soon be more competitive in the global marketplace. It is generally agreed upon by the majority of corporate managers and investors that improved corporate governance coming with Clause 49 is crucial in bringing Indian capital markets and governance standards up to par with respect to the rest of the world.

Clause 49 was created in 2000 and strengthened in 2004 by the Securities Exchange Board of India (SEBI) to ensure proper corporate governance of Indian companies. It applies to any company listed on Indian stock exchanges. Drivers for the reforms include the U.S. Sarbanes-Oxley Act and the Harshad Mehta and Ketan Parikh scams. Clause 49 establishes the following corporate governance guidelines:

- Establishes the minimum number of independent directors.
- Requires audit and shareholders' grievance committees.
- Requires a Management's Discussion and Analysis (MD&A) section.
- Requires a report on corporate governance in company annual reports.
- Requires disclosures of fees paid to nonexecutive directors.
- Limits the number of committees on which a director can serve.
- Requires CEO and CFO to certify financial results annually.

Southeast Asia[14]

East and Southeast Asian countries have a ways to go to reach Western levels of corporate governance. Exhibit 7.8 shows the World Bank's regulatory quality metrics for 2007 and 1996. With the exception of South Korea, nations in the region have made little progress in a decade. Taiwan, Malaysia, Thailand, Philippines, China, Indonesia, and North Korea have all lost ground.

Exhibit 7.9 shows the World Bank's control of corruption metrics for East and Southeast Asia. As with regulatory quality, there has been little progress over the last decade, and many countries in the region have lost ground.

The status of corporate governance can be summarized as follows:

- Companies are generally closely and/or family held. With the exception of Malaysia, diffused company ownership is relatively rare.
- CEO positions are held by nonprofessional managers in a majority of companies.
- Minority shareholders do not enjoy the rights of their Western counterparts.
- Disclosure and transparency is lacking, especially when there are conflict-of-interest issues.

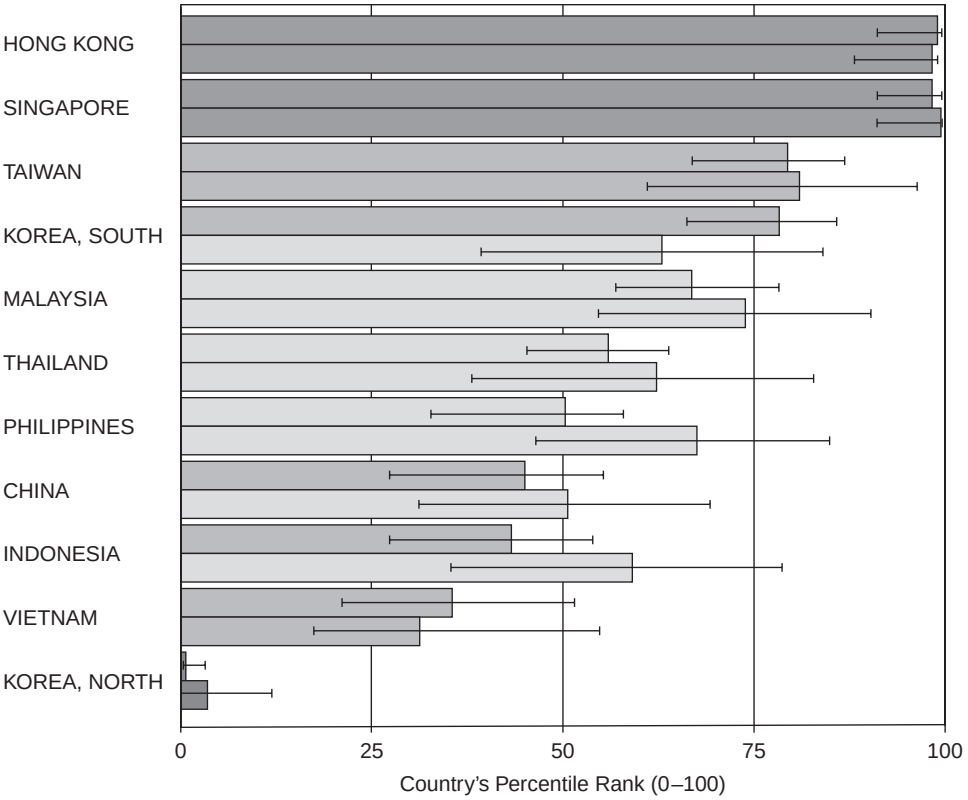


EXHIBIT 7.8 World Bank, Regulatory Quality: 2007 and 1996 (Top-Bottom Order)

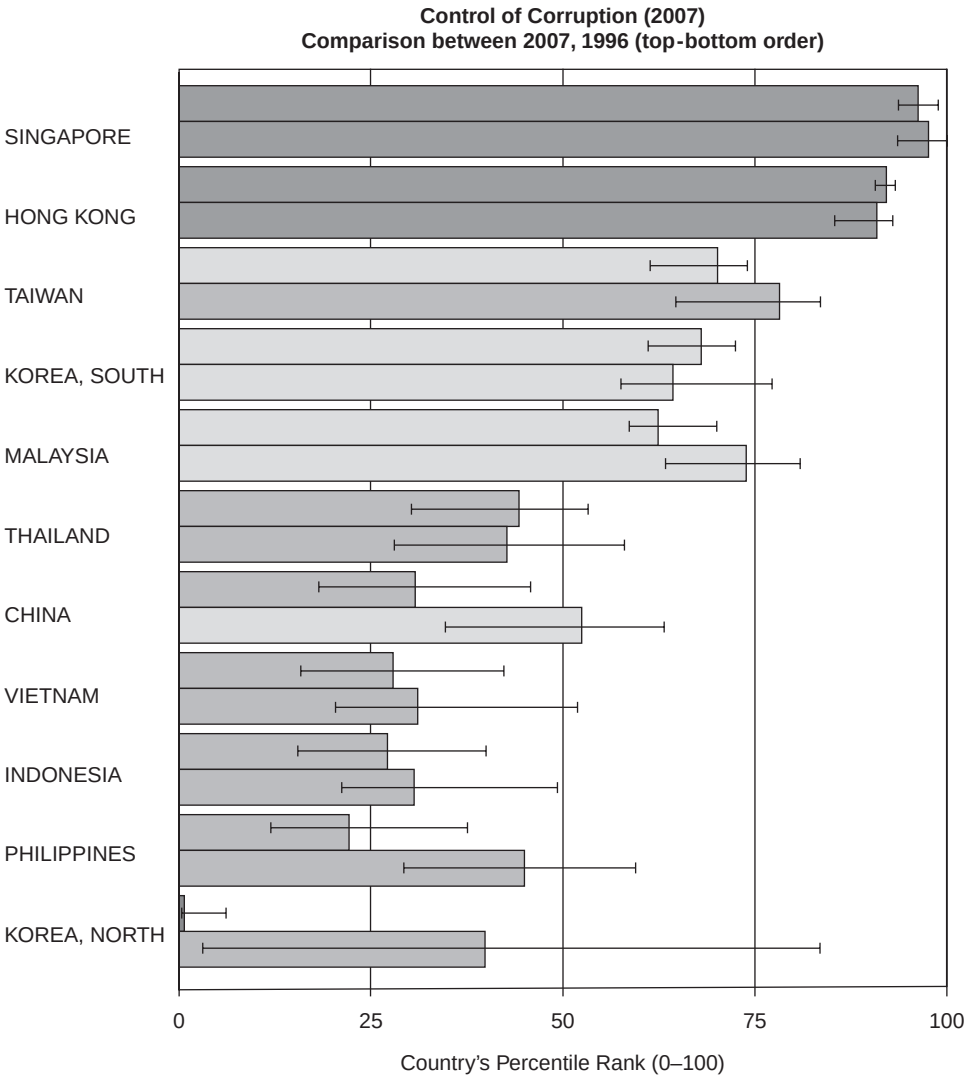


EXHIBIT 7.9 World Bank, Regulatory Quality: 2007 and 1996 (Top-Bottom Order)

- Independent directors are underrepresented as compared to their Western counterparts.
- Board expertise, proactive involvement, and committee involvement lag behind Western boards.

In spite of embracing the Organisation for Economic Co-operation and Development (OECD) Principles, and their desire to attract global capital, Southeast Asia has made little progress in closing the corporate governance gap with the Western economies. Only Hong Kong and Singapore enjoy Western governance levels for all six categories of governance tracked by the World Bank. Company-level governance

typically lacks Western counterparts in most or all critical areas. Auditors and regulators are not expert or motivated in evaluating the accuracy and transparency of financial reporting. The news media are not expert and/or free enough to provide the needed airing of corporate wrongdoing and regulatory weaknesses. These factors continue to present risks to investors and trading partners of these economies.

CONCLUSION

Investors and trading partners have many opportunities in their dealings with Asia's leading economies. The economic growth rates are the envy of the world. Before the current crash, Asian stock markets enjoyed very strong growth rates. Manufacturing expertise has risen to rival Western counterparts. India and South Asian countries are considered a very reliable means to outsource even the most complex information technology requirements.

The risks come in a variety of forms. Financial transparency accountability is an alien notion in the region. Companies are much more closely held than in the West, and boards are typically anemic. Regulators and auditors lack the expertise, motivation, and authority to compel significant improvements. As a result, investors and trading partners will need to continue to balance opportunities and risks. We have provided guidelines to partially mitigate the risks, but significant risk reduction will only come with major regulatory, accounting, and risk management reforms.

NOTES

1. Anthony Tarantino, *Governance, Risk, and Compliance Handbook* (Hoboken, NJ: John Wiley & Sons, 2008), 25–28.
2. "United States Public Debt," Wikipedia, accessed November 2008, http://en.wikipedia.org/wiki/United_States_public_debt.
3. "Economy of the People's Republic of China," Wikipedia accessed November 2008, http://en.wikipedia.org/wiki/Economy_of_the_People%27s_Republic_of_China#Systemic_problems.
4. "Economy of India," Wikipedia, accessed November 2008, http://en.wikipedia.org/wiki/Economy_of_India#cite_note-97.
5. "IBM India," Wikipedia, accessed November 2008, http://en.wikipedia.org/wiki/IBM_India#cite_note-bw-2.
6. Freedom House, Map of Press Freedom, www.freedomhouse.org/template.cfm?page=251&year=2007.
7. Central Intelligence Agency World Factbook, accessed November 2008), www.cia.gov/library/publications/the-world-factbook/rankorder/2078rank.html.
8. SupplyChainer.com, "Managing Supply Chain Risks in Asia," September 13, 2007, (Accessed November 2008), www.supplychainer.com/50226711/managing_supply_chain_risks_in_asia.php.
9. Vincent Kolo, "China's food contamination crisis deepens," *Chinaworker*, November 3, 2008, http://chinaworker.info/en/content/news/543/.
10. For a more detailed discussion of Japanese governance, see Anthony Tarantino, "Corporate Governance in China," Chapter 53, *Governance, Risk, and Compliance Handbook*.

11. Deloitte and Touche, "Basic Standard for Enterprise Internal Control helps raise governance and competitiveness of Chinese companies," July 2, 2008, www.deloitte.com/dtt/press_release/0,1014,sid%253D7062%2526cid%253D214900,00.html.
12. For a more detailed discussion of Japanese corporate governance, see Kouji Yamamoto, "The Guide to Global Compliance: The National Chapter—Japan," Chapter 59, in Tarantino, *Governance, Risk, and Compliance Handbook*.
13. For a more detailed discussion of Indian corporate governance, see Sanjay Anand, "The Current and Future States of Corporate Governance Culture and Regulation in India," Chapter 56, in Tarantino, *Governance, Risk, and Compliance Handbook*.
14. For a more detailed discussion of Southeast Asian corporate governance, see Lawrence Wasserman, "Southeast Asia Corporate Governance," Chapter 48, in Tarantino, *Governance, Risk, and Compliance Handbook*.

CHAPTER 8

Doing Business in Latin America: Lessons Learned and Best Practices for the Protection of Foreign Investors

Claudio Schuster and Pedro Fabiano

INTRODUCTION

The emerging markets present an excellent opportunity for global organizations to expand business. However, financial and regulatory risks are not always properly assessed when entering these markets.

During the 1990s, as a consequence of the globalization impulse in the economy, several companies started to explore new markets and possibilities. Many of them did not enjoy the global experience of traditional global companies such as Citigroup, General Electric, or General Motors. Organizations that have recently started their international endeavors often have never ventured beyond their national borders. They faced many new challenges, because in almost all the cases they lacked the proper experience to operate in countries with different legal and cultural traditions. They also did not have the proper international management, global networks, and the support and control systems, that the companies with decades of international experience had developed.

Latin America was one of these emerging market regions explored during the 1990s. It is a diverse region, and although many initiatives for integration were developed during that time, a great deal of differences exist in economic and political direction, political systems, economic integration, homogeneous rules and regulations in terms of commercial interchange, and important areas such as energy and telecommunications.

Traditional methods for project evaluation are not always applicable to the countries within the region. In the *discounted cash flow* (DCF) methodology, a company or financial asset is valued on the *time value of money* (TVM). With the TVM approach, all future cash flows are calculated and then discounted to give them a net present value. Typically, the appropriate cost of capital is used to determine the discount rate and may include a risk factor. DCF is well accepted in investment finance and corporate financial management.

The DCF methodology using the TVM concept considers the utilization of a cost-of-capital rate adjusted for inherent project risk and inherent country risk where the project is going to be implemented. A common practice is to consider the rate differential within the price of two sovereign bonds. In general, it is comparable to the local sovereign bonds price with U.S. Treasury bonds. Although this is a generally accepted measurement of risk in one moment of time, considering that the prices of the bonds are reflecting all the available information about the country, it does not consider the evolution of that risk during the time of the project.

Economic cycles in Latin America have been very short, and there have been significant changes in economic policies. Changes in the business rules, judicial insecurity, changes in the enforcement of laws, lack of protection for foreign investors, lack of transparency in government decisions, generalized corruption, radical and substantial changes in the economic path of the country, and political and social conflicts are only some of the issues that a foreign investor must face.

Investors also need to consider other criteria in order to evaluate investment decisions. The repayment period deserves special consideration. The longer this period is, the higher the risk to the project and its rate of return. A project that takes 15 years of cash flows in order to be profitable could not be completed in its lifetime. As a consequence of this, it is important to evaluate the time horizon when the project at least recovers the original investment.

Another important criterion to be considered is the divestiture flexibility of the project. It is not the same as the investment in a chain of commercial stores, with the possibility of easily closing the business and leaving the country, compared with the building of an automotive plant. The latter would require additional cash flows over the time in order to compensate the invested funds.

Exhibit 8.1 illustrates the more common risks a foreign investor could face in Latin America. It classifies examples of the different risks according to the criteria used by the World Bank in the study "Governance Matters V." Although these are not the only existing risks, they present a good example of what type of environment a foreign investor could find.

THE WORLD BANK INDICATORS

The study "Governance Matters V," published by the World Bank, presents a set of estimates of six dimensions of governance covering 195 countries and territories for the period 1996 to 2005. These indicators are based on several hundred variables measuring perceptions of governance, drawn from 25 separate data sources constructed by 18 different organizations. The individual measures of governance perceptions were assigned to six categories capturing key dimensions of governance. Six aggregate governance indicators were constructed, motivated by the broad definition of *governance* as the traditions and institutions by which authority in a country is exercised.

The first two governance clusters are intended to capture the first part of the definition of governance—the process by which those in authority are selected and replaced. The next two clusters summarize various indicators of the ability of the government to formulate and implement sound policies. The last two clusters summarize, in broad terms, the respect of citizens and the state for the institutions that govern their interactions.

This analysis does not necessarily mean that good business opportunities do not exist in the region. First, best practices for project analysis and evaluation should be adopted. In order to properly asses return possibilities and risks involved, it is important to invest in best practices for investor protection. It is also important to have the right mechanisms implemented that allow the investor to retire with the

Voice and Accountability *Voice and Accountability* includes a number of indicators measuring various aspects of the political process, civil liberties, and political rights. These indicators measure the extent to which citizens of a country are able to participate in the selection of government. This category also includes indicators measuring the independence of the media, which serves an important role in monitoring those in authority and holding them accountable for their actions.	Press manipulation with the risk of being demonized
Political Stability and Absence of Violence The second governance cluster is labeled *Political Stability and Absence of Violence*. This index combines several indicators which measure perceptions of the likelihood that the government in power will be destabilized or overthrown by possibly unconstitutional and/or violent means, including domestic violence and terrorism. This index also captures the idea that the quality of governance in a country is compromised by the likelihood of wrenching changes in government, which not only has a direct effect on the continuity of policies, but also, at a deeper level, undermines the ability of all citizens to peacefully select and replace those in power.	- Political instability - Increment of social conflicts - Permanent shocks and crisis with uncertain outcome - Constant change in global and international loyalties
Government Effectiveness *Government Effectiveness* combines into a single grouping responses on the quality of public service provision, the quality of the bureaucracy, the competence of civil servants, the independence of the civil service from political pressures, and the credibility of the government's commitment to policies. The main focus of this index is on "inputs" required for the government to be able to produce and implement good policies and deliver public goods.	- Low professional and moral level of the public servants - Delays in law enforcement and excessive level of lobby - Lack of infrastructure (energy, telecommunications, roads) - Personal hazard (guerillas, terrorism) - Growth in the impoverishment and analphabetism, with the domain of political machineries and radical and violent alternatives

EXHIBIT 8.1 World Bank Six Elements of Governance

Regulatory Quality The second cluster, *Regulatory Quality*, is more focused on the policies themselves. It includes measures of the incidence of market-unfriendly policies such as price controls or inadequate bank supervision, as well as perceptions of the burdens imposed by excessive regulation in areas such as foreign trade and business development.	- Limitations for funds transfers across the boards - Limitations of prohibitions to remit dividends to the home country - Lack of coherent and integrated regulations - Lack of economic integration - Changes in the economic orientation in short periods of time - Constant change of rules - Excessive bureaucracy - Limitations for exports - Limitations of raw material or capital goods imports - Difficulties accessing domestic or external financing - Nationalization and restatization of companies - Short and alternate periods of liberal economy vis-à-vis intervention
Rule of Law *Rule of Law* includes several indicators which measure the extent to which people have confidence in and abide by the rules of society. These include perceptions of the incidence of crime, the effectiveness and predictability of the judiciary, and the enforceability of contracts. Together, these indicators measure the success of a society in developing an environment in which fair and predictable rules form the basis for economic and social interactions, and importantly, the extent to which property rights are protected.	- Legal insecurity - High tax evasion - Lack of ethical conscience honoring contracts and commitments - Lack of knowledge of low importance to international regulations
Control of Corruption The final cluster, *Control of Corruption*, measures perceptions of corruption, conventionally defined as the exercise of public power for private gain. Despite this straightforward focus, the particular aspect of corruption measured by the various sources differs somewhat, ranging from the frequency of "additional payments to get things done," to the effects of corruption on the business environment, to measuring "grand corruption" in the political arena or in the tendency of elite forms to engage in "state capture." The presence of corruption is often a manifestation of a lack of respect of both the corrupter (typically a private citizen or firm) and the corrupted (typically a public official or politician) for the rules that govern their interactions, and hence represents a failure of governance according to the definition adopted by this study.	- Generalized corruption in the public and private areas - Excessive lobbying with an important degree of "friends of power"

EXHIBIT 8.1 (*Continued*)

major portion of profits are realized. These proper mechanisms are our major focus here—explaining the best practices in order to protect the interests of investors, which is in two forms, debt and equity.

PROTECTION OF DEBT INVESTORS

In the globalized world of today, it is necessary to have homogenous best practices for credit activities and to protect creditor rights. Due to the high level of volatility and the financial crisis of the 1990s, it is essential to integrate and standardize international practices and strengthen the international financial architecture. The World Bank (WB) and the International Monetary Fund (IMF) have developed a group of standards called *Principles and Guidelines for Effective Insolvency and Creditor Rights System*. These standards and codes are designed to evaluate and improve legal systems for credit matters, including access to credit facilities, protection mechanisms, risk management, workout procedures, commercial insolvency procedures, and related institutional and regulatory frameworks. The WB has also established the Reports on the Observance of Standards and Codes (ROSC), designed to evaluate the level of compliance of the country with the international standards and codes, and eventually recommend improvements.

The Principles and Guidelines for Effective Insolvency and Creditor Rights Systems cover specifically four areas:

1. Creditor Rights
2. Risk Management and Corporate Workouts
3. Commercial Insolvency
4. Institutional and Regulatory Frameworks

It is of the utmost importance for a creditor interested in Latin American countries to review the ROSC and evaluate the compliance of the country with the principles. This framework is essential when the risks mentioned in the introduction materialize. It is also important to assess the level of compliance with standards and codes, and the ability of creditors to exercise and protect their rights depending on the amount of economic value they would obtain from the transaction at the end of the day. The legal framework in terms of creditor rights and laws related to restructurings and bankruptcies will be fundamental at this stage.

The following case study exemplifies a typical situation of financing in emerging markets, and includes best practices in order to protect creditor rights. Although the case includes a good number of practices, the complete reading of the *Principles* would be beneficial for the reader.

LMP Case Study

Year 1992: LMP Co. (Lend Me Please Company, NYSE: LMPS) develops its activity in the energy industry, in the areas of production, transportation, and distribution of hydrocarbons and subproducts, in the internal market, and exports part of the production abroad. Due to increases in international demand and the prices for these commodities, the potential for new business has grown accordingly. Although the

activity is regulated inside the country, the political and economical orientation of the government is liberal, encouraging private activities. In this sense, the government has enforced laws toward monetary stability and protection of local and foreign investors, signed international treaties to promote foreign investments, and instituted measures for the development of local capital markets.

The majority of the stock holdings of LMP Co. are in the hands of local and international energy companies. The company is listed on the NYSE and the local stock exchange.

With this favorable environment and an encouraging outlook, LMP Co. has analyzed the expansion of its activities to a regional level. This will require a great degree of investments. The board and management of LMP Co. are analyzing the capital structure. The debt-to-equity ratio is around 30 percent, with some room for additional debt. The international financial market is in an expansion period, with an important level of liquidity and some appetite for investments in emerging markets with substantial returns. Considering this environment, the board and management of LMP Co. decided to be more aggressive, increasing the level of equity with the issuance of new stocks and the incorporation of a new international partner, and to issue debt up to a 50 percent level of the debt-to-equity ratio. The debt to be issued will be in the form of U.S. corporate bonds, multilateral agency facilities, and syndicated loans from commercial banks. Considering the magnitude of the project and the realization of earnings through the years, the lifetime of the debt also is going to be substantial, with a minimum of three years, a maximum of 10, and an average of seven years.

The financial strategy sounds aggressive, but the potential income justifies the decision. That is the way the creditors understand the deal. However, in the credit contracts, they include clauses for protection of their stake. Through the analysis of the project and the potential of the economy, the creditors are convinced that is a good deal and, following international best practices, they look for protection in front of potential contingencies.

First of all, the group of creditors analyzes the legal structure of the country and the mechanisms toward protecting credit and minimizing nonperformance and default. They review the ROSC in order to have a clear idea of the level of compliance of the country with international standards and codes. They also need to evaluate the existence of reliable procedures that enable credit providers and investors to more effectively assess, manage, and resolve default risk and to promptly respond to a state of financial distress of an enterprise borrower. They analyze the different mechanisms in the local laws and common practices that ensure transparency and celerity in the execution of their credits. They will also evaluate the need of inclusion of mortgages or other types of rights over the borrower assets, such as pledge over exports. In their contracts they will include covenants for the limitation of certain activities such as the level of indebtedness, payment of dividends, and some other important ratios. Finally, they will require the existence of proper documentation and records and properly audit over the activities and property of the company, including covenants that require timely information to the creditors of the company's activities. Last, they will evaluate the mechanisms and procedures that ensure efficient, transparent, and reliable methods for satisfying creditor rights by means of court proceedings or nonjudicial dispute resolution procedures.

Following the previous procedures, the group of creditors have evaluated the quality of the project from the economic and financial perspective, and covered their rights in case any credit risk materializes in the future.

Year 1998: The economy of the country is showing signs of deterioration, although it is far from a recession. Foreign investments have dropped significantly, reducing the economic growth rate. The economic indicators show signs of alert. The activity of LMP Co. has not suffered but will be impacted by increases in international hydrocarbons pricing. The management of LMP Co. analyzes the situation carefully because they have some new interesting projects that will require additional financing. The creditors are a little skeptical and hesitant to increase the level of financing. This is due to increases in the cost of credit for LMP Co., caused by increasing country risk, and the latest maturities of debt were refinanced and not canceled. They trust, however, in the economic future of the country, the attitude of the government, and health of LMP Co. A group of banks agreed to lend LMP Co. some additional credit facilities. The current debt/equity ratio of LMP Co. is 35 percent. However, the group of banks does not agree to increase this ratio over 45 percent, following international best practices and standards. They also make more strict default covenants, including a material adverse change (MAC) clause and a cross-default covenant. An MAC is a clause that triggers the total repayment of the facility in front of an extraordinary event, such as international financial crises. The cross-default covenant makes the credit required if any other credit of the company falls in default. LMP Co. accepts the inclusion of these clauses, and the increase of the cost of the credit itself, because they have no other viable alterative of financing.

Year 2001: The economic situation of the country is very bad. The ratios of the economy show the worst levels of the last 10 years. The country appears to be in a recession, with lagging investments and reductions in business activity. Social pressures are significant because poverty levels are high and unemployment is increasing. There are political crises almost every day with government agency heads resignations. The government, which is no longer pro liberal, decides to take some extreme economic measures. This includes a 200 percent devaluation of the currency, the freezing of internal tariffs, and limits on exports with imposition of higher taxes in the form of retentions.

During this period, LMP Co. has an important impact in its revenues. The company honors interest commitments with its creditors. However, without any chance of refinancing, the company decided to declare the default of its debt. The legal environment encourages the participants to understand and find solutions and agreements mutually beneficial, that allows the viability and continuity of the company.

The group of creditors is integrated by different types: individual bondholders, mutual funds, multilateral agencies, commercial banks, and others. In order to negotiate with the company, they established a credit committee (CC), and designate three of the most important creditors to sit on the committee. Local and international lawyers and investment banks were appointed as well as advisors.

The CC requires the availability of company information related with present activities and the company's financial situation. CC members discussed the possibility of a debt-to-equity swap, but some creditors are not will not agree with this proposal. The initial goal of the CC is to achieve a quick restructuring that enables the company's continuity in the future and the reinitiation of the payments to the

creditors as soon as possible. The creditors will need to be flexible and potentially accept reductions in interest rates or principal, and/or the extension in the payment schedule. Every effort is made to achieve an agreement that allows the continuity of the company as a going concern, and avoids liquidation or bankruptcy.

The CC knows that in an event where a negotiation and an agreement is not feasible, or the result of the conversations not acceptable for them, they always have the alternative of the judicial means in order to protect their rights. This analysis was made at the very beginning of the relationship with the company and the country, before lending the credit facilities to the company, with the objective of being covered and protected in front of these types of contingencies.

Finally, and after long negotiations, the CC agreed with LMP Co. on a new payment schedule with a reduction in interest rates, according to current international market conditions. With a major recovery in the economy, the company was able to honor its commitments without any problems. In the event of new contingencies in the future, the creditors of LMP Co. will have their risks covered in front of a workout, a restructuring, or legal actions.

Lessons Learned

1. The general political and/or economic environment can change dramatically in a short period in Latin America.
2. It is necessary to apply sound standards and codes and to be prepared for any contingency, in spite of the quality of the company.
3. Creditors need to be prepared to face changes and to be patient in order to recover their stake. These are the requirements in order to operate in a risky environment.

PROTECTION OF MINORITY OWNERS

According to the OECD Principles, the corporate governance framework should ensure the equitable treatment of all shareholders, including minority and foreign shareholders. All shareholders should have the opportunity to obtain effective redress for violation of their rights.

A poor or not properly enforced protection framework exposes minority owners to misuse or misappropriation by corporate managers, board members, or controlling shareholders. Also, corporate boards, managers, and controlling shareholders may have the opportunity to engage in activities that may advance their own interests at the expense of noncontrolling shareholders.

In providing protection to investors, a distinction can usefully be made between *ex ante* and *ex post* shareholder rights. *Ex ante* rights are, for example, preventative rights and qualified majorities for certain decisions. *Ex post* rights allow the seeking of compensations once rights have been violated.

In jurisdictions where the enforcement of the legal and regulatory framework is weak, some global companies have found it desirable to strengthen the *ex ante* rights of shareholders such as by low share ownership thresholds for placing items on the agenda of the shareholders meeting or by requiring a supermajority of shareholders for certain important decisions.

The World Bank research "Governance Matters V" shows that most countries in Latin America, including its first major economies, have low levels of regulatory quality, control of corruption, and law enforcement as compared to OECD countries where the Principles were developed.

As a response to these institutional limitations, the following case study is presented to illustrate a successful experience in which special risk management considerations and best practices were applied in protecting the interests of minority owners in Latin America.

LATCO Case Study

LATCO was a private company created as a result of the privatization of the electricity distribution business in a Latin American country and was organized under laws of this country. MGROUP, a local conglomerate dominated by a traditional family of the country owned 65 percent of LATCO. The minority owners were USCO, a U.S. financial institution listed on the New York Stock Exchange which owned 20 percent, and BRITCO, a British utilities company headquartered in London, which owned 15 percent. USCO and BRITCO designated the other two members of the board. HOLDCO, a holding company organized under the laws of state of New York, owned 100 percent of LATCO.

According to the shareholders' agreement signed in New York, LATCO was governed by the executive committee of the board. Members of the MGROUP family were appointed for the positions of chairman, vice chairman, and CEO. The shareholders' agreement also stated that MGROUP had the right to appoint the whole management team. However, USCO was entitled to designate the audit and compliance manager (ACM), who would report to the executive committee of the board.

During the first year of operations, the ACM had completed a comprehensive risk assessment, which revealed that LATCO was exposed to high risks of misdirected collections, undue payments, and unauthorized third-party related transactions. This risk assessment report also included recommendations for immediate corrective actions. The risk report was presented to the executive committee, and the CEO promised that he would effectively implement all the recommended actions to mitigate the risks. It was a commitment he would fail to keep.

Two years after the initial risk assessment, a follow-up report revealed that most of the recommendations had not been implemented. Management continued to report that the implementation of basic controls were "in progress." The CEO kept promising that he would fix the issues, and the executive committee, including the minority owner's representatives, accepted his excuses again and again. In addition, a special recommendation to investigate high-risk issues disclosed in the follow-up report was not considered by the minority owners.

The CEO influence and control over decisions grew dramatically during the first three years of operations. The minority owner's representatives followed him unconditionally. Moreover, the most important concern of the equity investors—the bottom line—was assured by the excellent profits of first three years. The company paid huge dividends, which kept everyone happy.

But the situation changed dramatically during the fourth year, when LATCO started to show liquidity problems and a significant decrease in reported profits. As a

result of this situation, the minority owners requested the ACM to perform an investigation. This investigation evidenced self-dealing, abuse by controlling shareholder, and unauthorized third-party-related transactions. Three main issues disclosed during the investigation were:

1. Cash withdrawals for $500,000 approved by the chairman of the board without supporting documentation. This transaction had not been recorded in the accounting system and was a clear violation of the shareholders' agreement. In fact, the shareholders' agreement required that any related-party transaction in excess of $100,000 should have been disclosed and approved by the minority owners.
2. Evidence obtained from public records confirmed that Consult LLC—a consulting firm that supposedly provided services to LATCO—was owned by LATCO CEO. This situation had not been disclosed to the partners. During the last year, LATCO had paid Consult $350,000 for services that were never rendered. Invoices were processed and paid with the written authorization of the CEO.
3. LATCO entered into a sale contract with an entity called Good Energy, which was a wholly owned MGROUP subsidiary. This 10-year deal, worth $100 million, included a discount of 26 percent over market price. Moreover, the review of public records showed that LATCO's CEO was president of Good Energy. It also found that the original payment terms had been changed from 7 to 70 days. This gracious amendment was signed by the LATCO chairman, a member of MGROUP. Once again, the shareholders' agreement had been violated because none of these decisions had been approved by the executive committee and no written justification was available.

The results of the investigation allowed the minority owners to obtain adequate and timely compensation from MGROUP, due to the application of the antifraud clauses of the shareholders' agreement.

Lessons Learned

This case study provides three critical lessons learned:

1. The low levels of the regulatory quality and enforcement of laws, regulations, and professional standards is commonplace in the region. As a result, the effectiveness of the GRC professionals depends heavily on a business decision and political support from the shareholders, regardless of the applicable legal, regulatory, and professional standards.
2. In this case, the ACM obtained full support for the investigation only because the minority owners had a business reason. The CEO was considered a "star," and the company was highly profitable. Everyone enjoyed remaining blissfully unaware of wrongdoing as long as business was good. Some businesspeople do not care about governance, risk, and compliance (GRC) until it is too late. This highlights the need to train business managers and other stakeholders about the different responsibilities of management, board members, external auditors, internal auditors, fraud examiners, compliance officers, and lawyers. A better

understanding of the GRC profession by those in senior positions will help to obtain their support in the protection of investors.

3. The recommendation to implement a code of ethics in LATCO was never discussed by the board. Fortunately, USCO lawyers had included some basic antifraud provisions in the shareholders' agreement. It would be a good practice for all companies to include antifraud provisions in their partnership and joint venture agreements. As shown in this case, the inclusion of control provisions in the shareholders' agreement was the only resource available to protect the interests of the minority owners. The participation of GRC professionals in the design of the shareholders' agreements is also highly recommended.

CONCLUSION

While this chapter is focused on Latin America, most all of these issues, lessons learned, and best practices transcend regions. The World Bank's metrics of corporate governance are always a good starting point in assessing the level of risk within any country—the lower the ratings, the greater the need for prudence in risk management.

The case studies demonstrate that environments with low standards of corporate governance, risk management, and regulatory compliance are inherently risky. In such conditions, financial transparency, robust risk management, whistle-blower protections, and minority investor rights are typically lacking. Governance and compliance are not baked into the DNA of many firms and are looked upon as merely a cost of doing business.

Unfortunately, the recent financial liquidity crisis that began in the United States and spread to Europe demonstrates once again that even countries receiving the highest World Bank ratings are not immune to scandals and catastrophic failures in risk management. Even more alarming, the current crisis occurred in the global financial service organizations under the most rigorous regulatory regimens and using the most sophisticated risk management processes and technologies.

Until the local and national environment compels a substantial upgrading of its governance and compliance standards, the best advice is to carefully evaluate each potential company investment to determine the integrity, risk management, and financial acumen of its executive leadership and board of directors. Even if the company is sound, national and regional conditions can change quickly and with little warning. Investors must be prepared for a wide variety of contingencies and be patient to recoup their investments.

CHAPTER 9

Mitigating Risk Exposure in Transitioning to the IFRS

Anthony Tarantino, Ph.D.

INTRODUCTION

The migration of the major economies from their local generally accepted accounting principles (GAAP) to the new International Financial Reporting Standards (IFRS) represents a much-needed modernization and standardization of accounting standards to accommodate a global economy. Without a global standard, how are investors to compare investment opportunities across dozens of disparate accounting standards? Under ideal IFRS conditions, an investor would be assured that companies across various national and regional jurisdictions, each reporting the same earnings per share, did in fact make the same amount of money. At present, this is far from a reality. Not only do the standards vary greatly, but the quality of accounting expertise, transparency, corporate governance, and oversight vary greatly.

The globalization of accounting standards under the IFRS is not without significant risks. Throwing out current practices that have been tried and proven in leading economies for something very new is never easy. Over the last century, the diversity in cultures and political and legal systems has created wide variations in accounting system and financial reports. Britain, France, Germany, and the Scandinavians all developed very strong but disparate accounting systems. With this realization and the need to promote commerce within the European Union (EU), Europe led the effort to develop an internationally accepted accounting standard.[1]

The accounting profession in the United States has a 75-year-old comfort level with U.S. GAAP. Japan's GAAP is modeled after U.S. GAAP and has been embraced for decades. As we detail in our *Governance, Risk, and Compliance Handbook*, the United Kingdom, Australia, and Canada have all enjoyed high levels of corporate governance using their existing accounting standards.[2] Each has embraced the IFRS. The United States has not scored as highly, but U.S. GAAP has not been targeted as a reason for major scandals of the last few decades. So it is not that the existing accounting systems were broken, but that the IFRS was a needed and welcome means to standardize standards around accepted best practices.

For most countries, there are no national equivalents to many of the new International Accounting Standards (IAS). There will be greater risk as they transition to areas unfamiliar to them. Less than 10 countries have followed at least some national

equivalents to the 31 IAS cited by Hussy and Ong in their *International Financial Reporting Standards Desk Reference*.[3]

It is no coincidence that those countries with the highest number of national equivalents to the IAS enjoy the highest corporate governance scores according to the World Bank. Taking an average of the World Bank's six elements of governance, Canada and Australia have achieved the highest ratings (92.3 percent and 91.5 percent, respectively) followed closely by Germany and the United Kingdom (both at 88.1 percent).[4] Australia, Canada, and the United Kingdom have all established over 20 national equivalents, but large numbers of national equivalent standards is no guarantee that the transition to the IFRS will be risk free. The United States has national equivalents to cover revenue recognition (Staff Accounting Bulletins [SAB] 101 and 104, detailed in our *Manager's Guide to Compliance*)[5] and stock options (SAB 107 and Sarbanes-Oxley Act [SOX] section 403, detailed in our *Governance, Risk, and Compliance Handbook*).[6] In spite of these, the United States has suffered major scandals in both areas.

In the United States, the principle-based IFRS will replace a very complex rules-based GAAP in which there are no U.S. equivalents for the following International Accounting Standards, according to Hussey and Ong[7]:

- IAS 2: Inventories
- IAS 10: Events after Balance Sheet
- IAS 11: Construction Contracts
- IAS 18: Revenue
- IAS 20: Accounting for Government Grants and Assistance
- IAS 28: Investment in Associates
- IAS 29: Accounting in Hyperinflationary Economies
- IAS 30: Disclosures in Financial Statements of Banks and Similar Financial Institutions
- IAS 31: Interest in Joint Ventures
- IAS 34: Interim Financial Reporting
- IAS 37: Provisions, Contingent Liabilities, and Contingent Assets
- IAS 39: Financial Instruments: Recognition and Measurement
- IAS 40: Investment Reporting
- IAS 41: Agriculture

Many of these international standards were published as early as 1993 but others are as recent as 2003, with limited adoption until January 1, 2005, when roughly 7,000 EU companies converted to the IFRS. Therefore, there is a major learning curve under way among accounting professionals.

Adding to the risk exposure is the lack of training among accounting and financial professionals in the IFRS. A survey by the American Accounting Association and KPMG, indicates that the first batch of U.S. accountants to be trained in the IFRS will not be available until 2011. The Securities and Exchange Commission's (SEC's) planned conversion is also 2011 and is unlikely to be delayed for fear of the United States losing more ground in its global competitiveness. In the survey, professors complain that their university administrations have not fully grasped the urgency to update curricula and texts.[8]

Educating the next generation of accounting professionals will be further complicated by the age of accounting professors—now averaging about 55 years. The American Institute of Certified Public Accountants (AICPA) Foundation is creating a doctoral program to address the coming shortfall, but this will do little to address the immediate need. It is hard to envision professors looking to retire in the next 5 to 10 years, looking forward to the major curricula changes IFRS will require.[9]

A cultural challenge in the transition will come in the critical role of international standard bodies under the IFRS. Unlike the United States and other legacy GAAPs, there is no single owner of the IFRS. This will require a transition from a national standards body to international standards body.

The U.S. embrace of international standards is fairly recent. In October 2002, the Financial Accounting Standards Board (FASB) and the International Accounting Standards Board (IASB) issued a letter of understanding known as the Norwalk Agreement, committing to the eventual convergence of U.S. GAAP with the IFRS.[10] The IFRS is now in use in many leading economies, including the EU countries, Hong Kong, Australia, Malaysia, Pakistan, India, Council of the Arab States of the Gulf (GCC) countries, Russia, South Africa, Singapore, and Turkey. As of mid-2008, more than 100 countries require or permit IFRS reporting. Of these, over 80 mandate the IFRS for all domestically listed companies.[11] Considering the global embrace of the IFRS since 2005, and that some EU firms have been able to use IAS since the late 1990s, the United States is 6 to 10 years behind the early adopters and major economies.

Potentially the largest cultural change from GAAP to IFRS is the requirement for greater judgment in preparing financial statements in that a very complex body of rules under GAAP is replaced with a much smaller body of principles—U.S. GAAP is roughly 10 times the length of the IFRS. Under U.S. GAAP, accounting professionals and auditors could rely on a checklist approach. Now they have to make judgments and disclosures as to why these judgments were made.[12]

To comprehend just how large of a cultural change the United States and other accounting professions face moving from a rules-based to a principles-based standard, consider the analogy of a simple traffic law covering driving a car through a four-way intersection. Using a rules-based approach there are many rules. Here are few of the most common:

- If a four-way stop sign, come to complete stop and yield to a driver on the right if both cars stop at the same time.
- If no stop sign, slow down to 15 miles per hour and visually check both ways.
- If a flashing yellow light, slow down and proceed when safe.
- If a flashing red light, stop and then proceed when safe.
- If controlled by a green/yellow/red light, proceed if green.
- If red, slow down and stop.
- If yellow, only proceed if you can get through the intersection before the light turns red.
- If an emergency vehicle is entering the intersection, do not cross the intersection.
- If there is no stop sign, slow to 15 miles per hour and only proceed when safe.

There are several others covering school buses, weather conditions, and so on.

Now consider under a principles-based approach, there is only one guideline—only go through the intersection when it is safe to do so. So ten or so specific rules with years of case law findings is replaced with a subjective guideline that is open to interpretation and requires an explanation of how the general guideline is being interpreted. The same one principle guidance also applies to railroad crossing, three-way intersections, and so on. So in total, dozens of specific rules would be replaced by one principle-based guideline. Imagine the court challenges to traffic citations under the one principle as litigants argue what "safe" really means—a lawyer's delight.

On its most basic level, accounting is the language of business. In the United States and many other economies, GAAP is the only language many accounting and tax professionals have ever known. The new mandated language is IFRS. This is akin to moving to a new country in mid life and having to give up your native language for an alien language. Clearly things will be lost in the translation.

What follows are some of the major areas of risk in the transition to the IFRS.

REVENUE RECOGNITION RISKS (IAS 18)

Revenue recognition has historically been a major cause of fraud and material weaknesses. According to the Committee of Sponsoring Organizations (COSO), half of all corporate fraud relates to revenue issues and one of the largest causes of material weaknesses and financial restatements under the Sarbanes-Oxley Act of 2002 (SOX).

The Securities and Exchange Commission (SEC) issued Staff Accounting Bulletin (SAB) 101 in 1999 and SAB 104 in 2003 to provide guidance to auditors and public companies on recognizing, presenting, and disclosing revenue in financial statements. Together SAB 101 and 104 describe the criteria for revenue recognition based on traditional accounting rules—revenue cannot be recognized until it is realized or realizable and earned. Under the U.S. SABs the following criteria must be met before revenue is recognized:

- There is persuasive evidence of an arrangement.
- Delivery has occurred or services have been rendered.
- The seller's price to the buyer is fixed or determinable.
- Collectability is reasonably assured.[13]

Common issues that have arisen from the transition to the IFRS include determining when transactions with multiple deliverables should be separated into individual components and with the manner in which revenue is allocated to the different components. Typically, U.S. GAAP emphasizes the separation and criteria for allocation, whereas IFRS emphasizes the transactions' economic substance.

Another issue arises in the accounting for customer loyalty programs. The IFRS treats customer loyalty programs as multiple-element arrangements—consideration is allocated to goods or services while award credits are based on their fair value from the customer's perspective. Under U.S. GAAP many companies have used an incremental cost model, which is substantially different from the IFRS's multiple-element approach.

For service transactions, U.S. GAAP prohibits use of the percentage-of-completion method unless the transaction qualifies under a specified contract types.

Others typically fall under the proportional/performance model. IFRS requires the use of the percentage-of-completion method unless it is not possible to reliably predict completion.

Construction contracts are also treated differently under U.S. GAAP and IFRS. The IFRS does not allow the completed-contract method, which may accelerate revenue recognition.

Another difference involves construction contracts, because IFRS prohibits use of the completed-contract method. This may result in the acceleration of revenue recognition under IFRS (depending on the specific facts and circumstances).

According to a KPMG webcast in October 2008, the following are some of the major issues organizations need to address for revenue recognition in making the conversion to IFRS:

- Implications from single revenue recognition models, regardless of industry, may permit reengineering of related processes and controls.
- Comparability with IFRS competitors.
- Change in how to separate multiple element arrangements and how fair value is measured (new data may need to be accumulated).
- Trigger point for revenue recognition.
- Long-term contracts that bridge transition date.
- First-mover considerations.
- Sales and incentive compensation."[14]

Under IAS 18, the U.S. criteria will change as indicated in Exhibit 9.1.

DERIVATIVES (IAS 39) AND HEDGING RISKS

Graham Holt describes derivatives as contracts for such financial products as options, forwards, futures, and swaps. Derivatives have the following characteristics: "Its value changes in response to the change in a specified interest rate, financial instrument price, commodity price, foreign exchange rate, index of prices or rates, credit rating, credit index or other variable it requires no initial net investment or the investment is small it is settled at a future date."[15]

Many times derivative contracts are entered into with little costs and therefore, prior to IAS 39, were typically not recognized in financial statements. Under IAS 39, derivatives are captured at their fair value. Any changes in fair value are recognized as either a profit or loss or as reserves, depending on whether hedging is used. Graham adds, "Where the derivative is used to offset risk and certain hedging conditions are met, changes in fair value can be recognized separately in reserves."[16]

Some EU national leaders, such as former French president Jacques Chirac recognized the risks in IAS 39 fearing that it would add to excess volatility in company earnings and balance sheets and as such scare investors away.

Hedging models in U.S. GAAP and IFRS are fairly similar and contain a significant amount of implementation guidance, which is unusual for other areas of IFRS. In some areas IFRS is more restrictive (e.g., prohibits shortcuts in measuring hedge effectiveness), but in other areas U.S. GAAP is more restrictive (e.g., foreign currency hedging risk). See Exhibit 9.2.

EXHIBIT 9.1 Revenue Recognition under IFRS and U.S. GAAP

Revenue Recognition: General Highlights	
IFRS: IAS 18	**U.S. GAAP: SAB 104**
There are probable future economic benefits to the seller.	There is persuasive evidence of an arrangement between the seller and the buyer. Collectability by the seller is reasonably assured.
Revenue can be measured reliably by the seller. Costs can be measured reliably by the seller.	The price is fixed or is determinable.
Significant risks and rewards of ownership are transferred to the buyer. The seller does not retain managerial involvement to the point of ownership nor retain effective control over the goods and services.	Delivery occurred and/or services have been rendered.
Revenue Recognition in Multiple Element Arrangements: Highlights*	
IFRS—IAS 18	**U.S. GAAP—EIFT 0021 and SOP 97-2**
There is limited guidance on separate components in a multiple element transaction and how fair value should be measured. Typically, separate arrangement consideration is based on the relative fair value (RFV) of separately identifiable components of a transaction.	Total consideration is allocated to each element based on FRV or the residual method (reverse residual method is not permitted). Substantial guidance is available on what constitutes fair value. Contingent revenue generally cannot be allocated to delivered elements.

*KPMG IFRS Institute webcast.

EXHIBIT 9.2 Derivative and Hedging Highlights under IFRS and U.S. GAAP

Derivatives Highlights	
IFRS	**U.S. GAAP**
Derivatives captured on financial statements at fair market value.	Derivatives not always captured on financial statements.
Hedging Highlights	
IFRS	**U.S. GAAP**
Hedging models are similar to U.S. GAAP. Does not permit the shortcut method—requires hedge effectiveness be tested and any ineffectiveness be captured as profit or loss. More restrictive than U.S. GAAP in the nature, frequency, and methods of assessing and measuring their effectiveness.	Hedging models are similar to IFRS. Permits in some cases, the shortcut method—bypassing effectiveness testing.

SHARE-BASED COMPENSATION AND PENSION RISKS

Companies that issue awards that vest ratably over time (e.g., 25 percent per year over a four-year period) may encounter accelerated expense recognition as well as a different total value to be expensed, for a given award, under IFRS.

Income tax expense (benefit) related to share-based payments may be more variable under IFRS. There are differences as to when an award is classified as a liability or as a component of equity. Those differences can have profound consequences, since awards classified as liabilities require ongoing valuation adjustments through each earnings reporting period. This can lead to greater earnings volatility.

There are a many differences between U.S. GAAP and IFRS in accounting for employee benefits. Some will result in more volatility, some will increase earnings, but others will decrease earnings.

Pension plans may see reduced volatility in that actuarial gains or losses would be recorded only within an IFRS equivalent of other comprehensive income, whereas U.S. GAAP allows asset values to be calculated to cover market movements up to five years.

IFRS does not require the presentation of various pension plan components as a net amount. This permits an organization to record the interest expense and return on plan assets of pension expense as part of their financing costs within an income statement.

In the *Governance, Risk, and Compliance Handbook*, we argue that there are better vehicles than share-based compensation. The major scandal in the United States showed the potential for abuse. It took section 403 of the Sarbanes-Oxley Act to end the abuse. A more fundamental issue is that compensation should be tied to activities over which an employee has some controls. Even at the executive level, share-based compensation tends to overemphasize short-term stock hikes over the long-term growth and prosperity of an organization. This has led some executives to take unrealistic risks and shortcuts in good governance. There is also an obvious accounting and administrative overhead to share-based compensation programs, which will increase under the IFRS. See Exhibit 9.3.

Pension plans, as opposed to share-based compensation, should see less volatility under the IFRS, but there are significant differences that present potential risks in making the transition.

EXHIBIT 9.3 Share-Based Compensation under IFRS and U.S. GAAP

Expense Recognition: Share-Base Compensation Highlights	
IFRS	**U.S. GAAP**
Accelerated expense recognition of stock options with graded vesting (e.g., 20% per year over five years).	Slower rates of expense recognition.
Greater tax rate variability over the lifetime of share-based payment awards.	More tax rate stability than in the IFRS.

EXHIBIT 9.4 Nonfinancial Assets under IFRS and U.S. GAAP

Nonfinancial Assets Highlights	
IFRS	**U.S. GAAP**
Last-in, first-out inventory costing prohibited.	Last-in, first-out inventory costing permitted.
Development costs capitalized under certain situations.	Development costs typically expensed as incurred.

NONFINANCIAL ASSET RISKS

Under U.S. GAAP, organizations were free to use first-in, first-out (FIFO); last-in, first-out (LIFO); or a weighted average cost method of valuing inventory. Typically, the benefits of LIFO versus FIFO depend on whether prices are rising or falling. IFRS simplifies the process by prohibiting LIFO. U.S. organizations that have used LIFO must convert and may be see changes in their earnings.

It may be advisable for organizations to evaluate the costs of converting away from LIFO through a pro forma financial analysis. If the costs are not significant, they may want to convert ahead of the mandatory deadlines. See Exhibit 9.4.

OFF-BALANCE-SHEET RISKS (FINANCIAL ASSETS)

In the *Manager's Guide to Compliance*, we detailed the continued complexity and therefore greater risks for investors and regulators in the continued use and abuse of off-balance-sheet arrangements *after* the enactment of the Sarbanes-Oxley Act (SOX). Enron's collapse was caused by its flagrant abuse of special-purpose entities (SPEs). They were kept off its balance sheet to hide massive obligations that the company could not meet. The irony comes in that most SOX complaints are over the increased costs for improving internal controls under Section 404. Section 401 is designed to improve the control over off-balance-sheet arrangements, but did little to prevent large global banks from hiding their subprime mortgage exposure with their use.

In the case of Citigroup, investors were blind to massive risk exposure from the subprime mortgage market until their new CEO, Vikram Pandit, brought these off-balance-sheet obligations, known as qualified special-purpose entities (QSPEs) back on the balance sheet. Citigroup then promptly wrote them off—part of an $18 billion write-down in January 2008.

Under current U.S. GAAP, certain loans such as those linked to credit card debt and to high-risk mortgages, can be kept off balance sheet with the use of such QSPEs. With the IFRS, the primary goal is to increase controls using its principles-based approach. This makes it more difficult to design financial products in a way to keep them off a company's balance sheet.

Considering the abuse and increased risk, one can argue that limiting off-balance-sheet obligations will provide investors and regulators with better financial

transparency. While it may seem to hurt organizations by requiring greater capital reserves, reducing off-balance-sheet arrangements will have benefits in mitigating the types of risk taking that have caused the greatest financial crisis since the 1930s.

There is much at stake in the U.S. GAAP/IFRS convergence, and it is a hotly debated issue on both sides of the Atlantic. On September 15, 2008, the International Accounting Standards Board (IASB) met to discuss derecognition of financial assets (off balance sheet) under the U.S. FAS 140 and IFRS IAS 39.[17] Earlier in the year, the Financial Stability Forum suggested making off-balance-sheet an urgent priority because of the global financial crisis. It argued that improved standards were essential. The IAS board agreed in its September meeting that it is urgent to improve off-balance-sheet accounting and disclosure and to bring about a convergence with the U.S. GAAP. The IAS board noted that off-balance-sheet entities created an incorrect perception that there was no significant risk for organizations. The board argued that substantially reducing off-balance-sheet arrangements would provide a much clearer view of an organization's risks in its financial disclosures.

Off-balance-sheet arrangements in financial statements can arise as a result of derecognition standards in which assets are removed from balance sheets through securitizations, or through consolidations such as SPEs. There are significant differences in the treatment of off-balance-sheet arrangements between the IASB and the U.S. Financial Accounting Standards Board (FASB). The IASB and FASB are moving quickly to converge their off-balance-sheet standards with the goal that off-balance-sheet risks be clearly identified and presented in financial statements.

The debate about derecognition has been going on for years, but no satisfactory and lasting solution has been forthcoming. IAS 39 covers off-balance-sheet arrangements and is more restrictive in permitting their use than the comparable U.S. standard, FAS 140. But the proposed amendments to FAS 140 will make it more difficult to derecognize assets than in the past.

While there are very legitimate uses for off-balance-sheet arrangements, their abuse was behind the Enron scandal, one of the greatest accounting frauds in history that destroyed Arthur Andersen and sparked SOX. They continue to be abused during the current financial crisis by banks hiding high-risk obligations until forced to bring them back on their balance sheet. Maybe the best advice to organizations facing the IFRS is to take more conservative approach by weaning themselves from off-balance-sheet arrangements sooner than later. See Exhibit 9.5.

EXHIBIT 9.5 Off-Balance-Sheet Arrangements under IFRS and U.S. GAAP

Financial Assets Highlights: Off-Balance-Sheet Arrangements	
IFRS	**U.S. GAAP**
Greater restrictions over off-balance-sheet arrangements—requiring partial or full balance recognition. In financial services this translates into greater capital requirements.	Greater use of off-balance-sheet arrangement permitted. In financial services, this reduces capital requirements under the Basel II Accords.

TAX LIABILITY RISKS

While U.S. GAAP and the IFRS share many tax principles, there are differences in how they calculate liabilities, contingencies, and deferred taxes. This will require adjustments in an organization's tax accounts.

A major difference is the tax expense impact of cross-border inventory transfers within a consolidated group. Under the IFRS, deferred taxes on intragroup profits are determined by reference to the buyer's tax rate. Under the U.S. GAAP, income tax effects resulting from intragroup profits are deferred at the seller's tax rate. This may translate into less volatility in financial statements under the IFRS.

Under IFRS, all future increases or decreases in equity-related deferred tax asset or liability accounts must be traced back to equity. Under U.S. GAAP, any subsequent changes arising from legal and tax rate changes in deferred taxes are recognized through an operations statement even in cases where the related deferred taxes initially arose in equity.

Under U.S. GAAP, any income tax effects resulting from intragroup profits are deferred at the seller's tax rate and recognized upon the sale to a third party. Under IFRS, deferred taxes must be captured based on the buyer's tax rate at the time of the initial transaction. See Exhibit 9.6.

OTHER LIABILITY RISKS

The IFRS treats provision accounting in a manner that may result in recognizing expenses earlier. This includes differing thresholds as to when provisions are to be established.

IFRS has a higher threshold for the recognition of contingent assets associated with insurance recoveries by requiring that they be virtually certain of realization. U.S. GAAP allows earlier recognition of contingent assets associated with insurance recoveries.

EXHIBIT 9.6 Tax Liability under IFRS and U.S. GAAP

Tax Liability Highlights	
IFRS	**U.S. GAAP**
Deferred taxes on intragroup profits are determined by reference to the buyer's tax rate.	Income tax effects resulting from intragroup profits are deferred at the seller's tax rate.
Future increases or decreases in equity-related deferred tax asset or liability accounts must be traced back to equity.	Subsequent changes arising from legal and tax rate changes in deferred taxes are recognized through an operations statement even if the related deferred taxes initially arose in equity.
Deferred taxes are captured based on the buyer's tax rate at the time of the initial transaction.	Income tax effects resulting from intragroup profits are deferred at the seller's tax rate and recognized upon the sale to a third party.

EXHIBIT 9.7 Other Liabilities under IFRS and U.S. GAAP

Other Liability Highlights	
IFRS	**U.S. GAAP**
Creates a higher threshold for the recognition of contingent assets associated with insurance recoveries—must be virtually certain of realization.	Allows earlier recognition of contingent assets associated with insurance recoveries.
Probable means "more likely than not." This is a criterion in liability recognition.	*Probable* means "as likely to occur." This is a criterion in liability recognition.

There are differences in how IFRS and U.S. GAAP interpret the term *probable*. Under IFRS, *probable* means "more likely than not." Under U.S. GAAP, *probable* means "as likely to occur." This is important because both frameworks use the term *probable* as a criterion in liability recognition. See Exhibit 9.7.

FINANCIAL LIABILITIES AND EQUITY RISKS

Under IFRS, warrants are treated as derivative instruments and therefore marked to market through earnings Under U.S. GAAP, warrants issued in the United States can be net share settled and as such classified as equity.

Under IFRS, more instruments are likely to be classified as liabilities, and not as equity.

The U.S. GAAP create a more narrow definition of what instruments constitute a liability than does the IFRS. Both U.S. GAAP and IFRS define financial liabilities and require that financing instruments be assessed to determine as to whether they are defined and treated as liabilities. Typically, financial instruments that do not meet the definition of a liability are classified as equity.

The IFRS has created one comprehensive standard, IAS 32, to determine the appropriate classification of an instrument as a liability or as an equity. Under IFRS, contingent settlement provisions and puttable instruments are more likely to result in liability classification. The goal of IAS 32 is to assess the substance of contractual arrangements, not their legal form. Unlike the IFRS, there is no one comprehensive guidance under U.S. GAAP. The guidance comes from SEC rules, Emerging Issues Task Force (EITF) issues, and FASB standards. See Exhibit 9.8.

BUSINESS COMBINATION RISKS (MERGERS AND ACQUISITIONS)

U.S. GAAP guidance is being updated to more closely match the IFRS. One of the most significant is the change requiring organizations to expense acquisition costs that had been capitalized under the old guidance. Another change requires that restructuring costs be recognized separately from a business combination after the combination is completed.

EXHIBIT 9.8 Financial Liabilities and Equity Highlights under IFRS and U.S. GAAP

Financial Liabilities and Equity Highlights	
IFRS	**U.S. GAAP**
Warrants of are treated as derivative instruments and therefore marked to market through earnings.	Warrants issued in the United States can be net share settled and as such classified as equity.
Instruments are more likely to be classified as liabilities, and not as equity.	Creates a more narrow definition of what instruments constitute a liability than does the IFRS.
One comprehensive standard, IAS 32, to determine the appropriate classification of an instrument as a liability or as an equity.	No one comprehensive guidance to determine the appropriate classification of an instrument as a liability or as an equity. The guidance comes from SEC rules, EITF issues, and FASB standards.

Even under the new guidance, major differences will remain in recognition at the date of acquisition as to how they impact contingent liabilities. In addition, there will be differences in the subsequent measurement of contingent liabilities that may result in more volatility under IFRS. The new U.S. FAS 141 makes significant changes in the treatment of acquisitions and noncontrolling interests in a subsidiary. It will also continue the movements toward improved financial disclosure and fair value financial reporting.

Under IFRS indicators of control are utilized, some of which individually determine the need to consolidate. In cases where control is not apparent, consolidation is based on an overarching assessment of all of the relevant facts. This includes the risk and benefit allocations between the two organizations. Under the IFRS, consolidation is required when one organization has the means to govern the financial and operating policies, procedures, and processes of another organization in order to obtain benefits.

U.S. GAAP uses a two-tiered model of consolidation—one focused on the organization's exposure to the risks and rewards from the other organization's activities (where the party that participates in the majority of the entity's economic impact consolidates such operations), and one focused on voting rights (where the investor owning more than 50 percent of an entity's voting interests consolidates the investee's operation).[18]

All organizations are evaluated to determine if they meet the requirements of a variable-interest entity (VIE). There are three requirements of VIEs: they are not self-supportive, they have variable interests in the VIE by providing it with financial support, and they must be the VIE's primary beneficiary, such as by absorbing more than half of expected losses or receiving more than half of expected residual returns.[19] If they are deemed a VIE, consolidation is based only on economic risks and rewards—decision-making authority plays no role in the consolidation decision.

Even with the major changes coming with FAS 141, the transition to the IFRS will present very significant changes and increased risks in the area of mergers and acquisitions. The IFRS approach is alien to U.S. organizations, and there will be

EXHIBIT 9.9 Business Combinations under IFRS and U.S. GAAP

Business Combination Highlights	
IFRS	**U.S. GAAP**
Indicators of controls are utilized to determine the need to consolidate.	Two-tiered model of consolidation is used—one focused on the organization's exposure to the risks and rewards from the other organization's activities, and one focused on voting rights.
When control is not apparent, consolidation is based on an overarching assessment of all relevant facts such as risk and benefit allocations between the two organizations.	If an organization is deemed to meet the requirements of a variable-interest entity (VIE). As a VIE, consolidation is based only on economic risks and rewards and not on decision-making authority.
Consolidation is required when one organization has the means to govern the financial and operating policies, procedures, and processes of another organization in order to obtain benefits.	

a steep learning curve. They will be challenged to make the changes to comply with FAS 141 and then have to change again within a few years to the IFRS. See Exhibit 9.9.

FINANCIAL SERVICES INDUSTRY RISKS

Quantification of risk management is a key requirement within the financial services industry. The Basel II accords in banking and the Solvency II accords in insurance mandate much greater enterprise risk management (ERM) than in the past. Operational risk management, mandated in both accords, requires massive amounts of historical loss data over extended periods of time. The IFRS requires more detailed disclosure and analysis as to how various risk scenarios are managed and how they impact cash flows. The catastrophic failures in risk management within the financial services industry will intensify the scrutiny under the IFRS.[20]

Many banks are looking to voluntarily opt in to the advanced measurement approach (AMA) under the Basel II accords or face substantially higher capital charges. Solvency II has similar requirements, has been adopted in the European Union, and is likely to become a global requirement for insurers in the future. The major rating agencies have already put insurers on notice that they need to adopt robust ERM, which includes the AMA, or face being downgraded. Insurers and bankers are going to be hard pressed to improve their controls, stress testing, risk monitoring infrastructure, and disclosure in order to comply with the IFRS.

Regulators and auditors will have little patience with laggards, especially with current financial crisis. An emotional overreaction swept the United States after the failure of Arthur Andersen, stemming from the Enron scandal. Such a response can be expected with even greater fervor now, given the extent of the damage.

Credit risk will also be impacted by the IFRS conversion. More stringent analysis and accounting requirements change the valuation of assets. This includes booking on the balance sheet the likely impairment of asset values based on market fluctuations. It also includes documenting the validation of the risk frameworks and technologies that an organization uses.

Finally, the IFRS will substantially change the use of off-balance-sheet arrangements which non-IFRS banks have used to lower their capital requirements.

CONCLUSION: SUGGESTIONS TO REDUCE THE CONVERSION RISKS

Converting U.S. organizations to the IFRS will increase their risk exposure in some very fundamental ways. Under Section 302 of SOX, the chief executive and chief financial officers must certify to the accuracy of their financial statements. For these executives, this has been a traumatic experience. Many executives have lost their jobs after being forced to declare material weaknesses or restate earnings (explained in detail in our *Manager's Guide to Compliance*).

Chief technology and information officers have not been immune to the trauma, as finance executives continue to pressure them to improve data access, quality, and standardization. Like financial executives, they have suffered much higher employee turnover rates than in the past and than their Japanese and European counterparts. The conversion to the IFRS will create even greater demands on information technology. This includes data capture, storage, accessibility, normalization, quality, analytics, and modeling. Fair value accounting, weighted averages, and extensible business reporting language (XBRL) will require more sophisticated and timely systems increase demands to calculate values.

The trauma from U.S. SOX is now leveling off, but will be replaced with the demands to meet the IFRS. Of course, the global financial crisis will dramatically increase pressure to provide timely and accurate financial information. Auditors, regulators, rating agencies, and shareholders will be unlikely to show any tolerance for mistakes during the transition to the IFRS.

Here are some recommendations to ease the pain of the IFRS transition for U.S. and other non-IFRS organizations:

- *Invest in training and upgrading your financial resources.* The good news is that there is a large body of publications and seminars from which to draw on.
- *Take a hard look at your current financial resources.* Change is always traumatic and exposes personnel weaknesses not obvious in maintaining the status quo. Financial professionals, who have been comfortable under GAAP, may be ill suited as change agents under IFRS.
- *Consider hiring financial resources from the EU with IFRS experience.* This may seem to be an expensive option, but the alternative is to pay greater consulting and auditing fees. It is better to have your own internal consultants who have a vested interest in your success.
- *Perform a classic gap analysis.* This should include ranking your financial statements as to the most significant changes under the IFRS. The areas of the greatest change will present the greatest risks, but can also present the greatest opportunities.

- *Create pro forma financials under IFRS.* If practical, create pro forma financial statements as the organization would look like under IFRS. If it makes financial sense and reduces risk, start converting some elements earlier than legally mandated. It may make sense to start reducing your more risky practices around revenue recognition, off-balance-sheet arrangements, combinations, and so on before legally mandated.

NOTES

1. Naomi S. Soderstrom and Kevin Jialin Sun, "IFRS Adoption and Accounting Quality: A Review, Social Science Review Network," October 2007, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1008416.
2. Anthony Tarantino, *Governance, Risk, and Compliance Handbook* (Hoboken, NJ: John Wiley & Sons, 2008).
3. Roger Hussey and Audra Ong, *International Financial Standards Desk Reference* (Hoboken, NJ: John Wiley & Sons, 2005).
4. Tarantino, *Governance, Risk, and Compliance Handbook*.
5. Tarantino, *The Manager's Guide to Compliance* (Hoboken, NJ: John Wiley & Sons, 2006), 125–134.
6. Tarantino, *Governance, Risk, and Compliance Handbook*, pp. 111–120.
7. Hussey and Ong, *International Financial Standards Desk Reference*.
8. Penny Sukharj, "First Batch of IFRS Graduates Only Ready in 2011," *Accountancy Age,* September 5, 2005.
9. APICA Foundation, "Doctoral Scholoars Program in Accounting Created by CPA Profession," July 30, 2008, www.ficpa.org/fs_ficpa/publicfiles/national_news/aicpa/2008/accountingdoctoralscholarships.pdf.
10. Hussey and Ong, *International Financial Standards Desk Reference,* p. 31.
11. Wikipedia, IFRS, http://en.wikipedia.org/wiki/International_Financial_Reporting_Standards.
12. See Timothy Flynn, chairman of KPMG International, Interview, "US Warming to IFRS as It Moves on from GAAP," *The Financial Times*, September 4, 2008.
13. See Tarantino, *The Manager's Guide to Compliance,* Chapter 17: "Revenue Recognition Requirements: U.S. SAB 101 and 104."
14. KPMG IFRS Institute webcast, "IFRS for Technology Companies: Closing the GAAP?" October 8, 2008, www.kpmgifrsinstitute.com/ContentDetails.aspx?content_id=2016.
15. Graham Holt, "IAS 39, Financial Instruments: Recognition and Measurement II," Association of Chartered Certified Accountants (ACCA)web site: www.accaglobal.com/members/publications/accounting_business/cpd/2806959.
16. Ibid.
17. International Accounting Standards Board Meeting, "Project: Derecognition of Financial Assets," London, September 15, 2008, www.iasb.org/NR/rdonlyres/19AF50E3-8F89-4E27-B19A-461F63230E0D/0/Derec0810b07obs.pdf.
18. Alan Reinstein, Gerald H. Lander, and Stephen Danese, "Consolidation of Variable-Interest Entities Applying the Provisions of FIN 46(R)," *CPA Journal,* August 2006, www.nysscpa.org/cpajournal/2006/806/essentials/p28.htm.
19. Ibid.
20. PriceWaterhouseCoopers (PWC), "IFRS and Risk Management: IFRS—Global Reporting Revolution," April 2004, www.pwc.com/images/gx/eng/fs/insu/0304ifrsrisk.pdf.

CHAPTER 10

Quantitative Operational Risk Management Methods

Deborah Cernauskas, Ph.D.

INTRODUCTION

Operational risk is one of numerous risks (see Exhibit 10.1) monitored, managed, and controlled by financial firms. Its importance has grown exponentially over time, in part due to the spectacular operational loss events such as the collapse of Barings Bank PLC. The Basel Committee on Banking Supervision (BCBS[1]) published papers titled "A Framework for Internal Control Systems in Banking Organizations" (1998) and "Sound Practices for the Management and Supervision of Operational Risk" (2003), which laid the foundation for measuring operational risk for financial institutions, but only from a capital perspective. Basel II defines operational risk as "the risk of losses resulting from inadequate or failed internal processes, people, and systems or from external events" (see Exhibit 10.2).

This definition includes legal risk, but excludes strategic and reputational risks. While this definition holds true for any industry, the specific operational risk events will vary from company to company. A manufacturing company will have many of the same operational risks as a bank (e.g., fraud and computer failures), but will also have industry-specific risks (e.g., hazardous materials handling and physical injuries).

Operational risk is an issue for all companies, but its scope is so vast that it is hard to define and equally hard to measure. Unlike market or credit risk, there is no standard unit of measure for operational risk, even within the same company. For example, the standard unit of market risk is the asset whose price change may cause a loss. Operational risk is too diverse to have a standard unit, and it is generally characterized as those risks related to business, crime, disaster, information technology (IT), and regulatory compliance, but excludes strategic processes and reputational risk. It is the hardest risk to anticipate and has the potential to be of devastating magnitude to the finances of the company. Although operational risk has always been an issue for firms, the quantification of operational risk has come to the forefront since Basel II's inclusion of a capital charge for operational risk.

Currently, many industry professionals and academics are struggling to identify and quantify the numerous risks that fall under the canopy of operational risk while there is a scarcity of data available. Modeling efforts to quantify operational risk will not be very successful until adequate internal and external data are available.

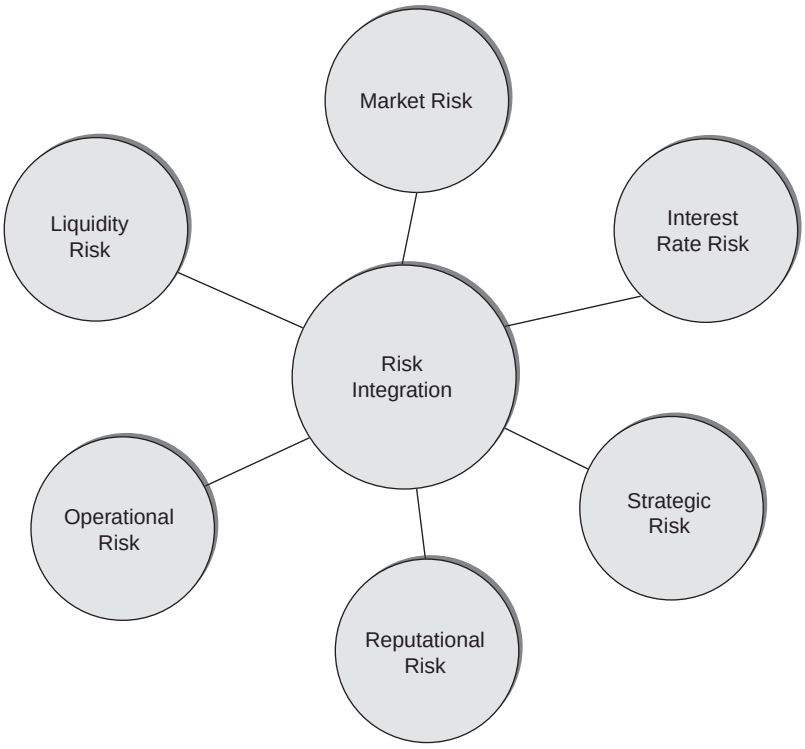


EXHIBIT 10.1 Risk Types

To this end, over the past couple of years, some companies and consortiums have been actively compiling loss databases. The Operational Risk Exchange is a loss data consortium of global banks formed to help the industry comply with Basel II capital regulations and to enhance their members' internal risk management efforts.

A further complicating factor in the measurement of risk is the interdependency of market, credit, and operation risk. Whereas Basel II treats these risks as independent, this is not always a good assumption. Consider a trading firm where operational risk and losses arising from human and technological errors can easily transform itself into market and credit risk. In 1995, Barings PLC declared bankruptcy due to

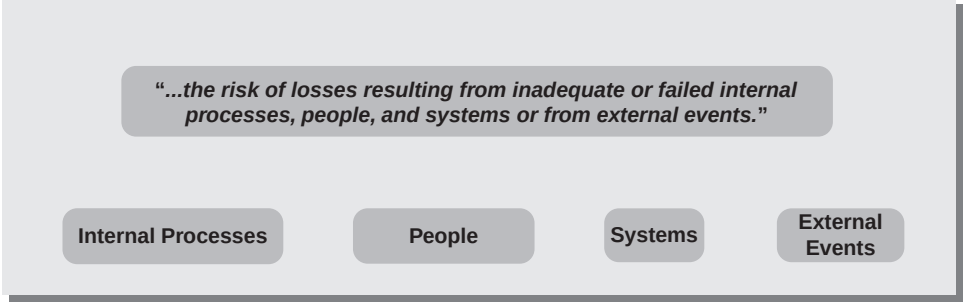


EXHIBIT 10.2 Operational Risk Definition

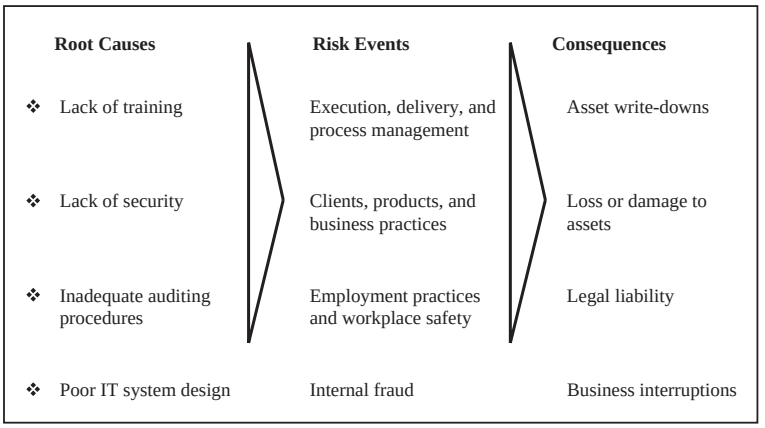


EXHIBIT 10.3 Operational Risk Categories

the actions of a single trader, who lost $1.3 billion dollars in derivatives trading. The derivatives market risk Barings succumbed to was due to a lack of proper controls, an operational risk.

Operational risk management is the process of identifying, measuring, or assessing operational risk and then developing strategies to manage/mitigate the risk. It spans root causes, events, and consequences (see Exhibit 10.3). Most large companies have operational risk management processes in place and they know that taking risks is part of doing business and that managing risks is critical to their success. In spite of this, there is still a huge gap in the area of operational risk management. The onus is placed on the individual functional areas to manage these risks as opposed to an enterprise approach. Almost all of the well-publicized corporate scandals can be attributed to failures in identifying and managing internal sources of risks. This gives corporate managers hope the high-dollar-loss scandals can be stopped with the implementation of the right governance, process monitoring, and process controls.

The focus of this chapter is on providing an overview of quantitative methods available to measure and manage risk exposures.

OPERATIONAL RISK OVERVIEW

Even though the Basel Committee addresses the banking industry, the underlying fundamentals can be applied to any industry or organization. Basel II believes that deregulation, globalization, and growing sophistication of financial technology are making the activities of banks and thus their risk profiles more complex. The same can be said of any type of business (e.g., manufacturing, mining, health care, food and drug, etc.) Some of the examples of growing sophistication cited by the Basel Committee are:

- Greater use of more highly automated technology has the potential to transform risks from manual processing errors to system failure risks.
- Growth of e-commerce brings with it potential risks that are not fully understood.

- Large-scale acquisitions, mergers, demergers, and consolidations test the viability of new or newly integrated systems.
- The emergence of banks acting as large volume service providers creates the need for continual maintenance of high-grade internal controls and backup systems.
- Banks may engage in risk mitigation techniques (e.g., collateral, credit derivatives, netting arrangements, and asset securitizations) to optimize their exposure to market risk and credit risk, but which in turn may produce other forms of risk (e.g., legal risk).
- Growing use of outsourcing arrangements and the participation in clearing and settlement systems can mitigate some risks, but can also present significant other risks to banks.

The term *operational risk* carries different meanings to different organizations. No matter how a particular organization defines operational risk, a clear understanding of what is meant is critical to the effective management and control of this risk. Any of the following events categorized as operational risks can result in substantial losses:

- Internal fraud (e.g., employee theft, insider trading, etc.).
- External fraud (e.g., robbery, forgery, check kiting, and computer hacking).
- Employment practices and workplace safety (e.g., workers compensation claims, violation of employee health and safety rules, organized labor activities, discrimination claims, and general liability).
- Clients, products, and business practices (e.g., fiduciary breaches, misuse of confidential customer information, improper activities on the bank's account, money laundering, and the sale of unauthorized products).
- Damage to physical assets (e.g., terrorism, vandalism, earthquakes, fires, and floods).
- Business disruption and system failures (e.g., hardware and software failures, telecommunication problems, and utility outages).
- Execution, delivery, and process management (e.g., data entry errors, collateral management failures, incomplete legal documentation, unapproved access given to client accounts, nonclient counterparty misperformance, and vendor disputes).

QUANTITATIVE METHODS

Quantitative analysis refers to the use of numerical and statistical techniques to gain insight and extract information from data. Quantitative analysis is data driven, and data is central to everything. As the saying goes: If you can't express something in the form of numbers, you really don't know much about it. If you don't know much about it, you can't control it. If you can't control it, you are at the mercy of chance and, hence, why bother with it.

Market, credit, and insurance risks rely heavily on statistical analysis of historical data for quantification. There is an enormous amount of research and available historical data in this space and a number of sophisticated tools are available for modeling complex scenarios to understand and mitigate risk. The same cannot be

said for measuring, modeling, and managing operational risk. It is not always easy to collect data on each and every business process within a company. Even if there are data being collected, it may not be in the desired format or may not meet the needs of quantitative analysis. As data are collected, the ability to measure risk exposures and develop monitoring capabilities based on risk analytics will expand tremendously.

MODELING APPROACH OPERATIONAL RISK

Much is being written about the failure of risk management to spot the dangers in the credit market despite the adoption of Basel II risk quantification. Due to current financial problems in the economy, model and quant bashing is in vogue. Fortunately, this attitude is not held by all. Tom Garside, global head of the finance and risk practice at Oliver Wyman in London, was recently quoted in an industry journal[2] that some banks had risk models that worked, giving advanced warning of a credit bubble and took action to reposition themselves in advance of the bursting bubble. In the same article, Aaron Brown, a risk manager at AQR Capital Management stated "... the system worked in every way, except nobody paid attention. People just didn't trust the models."

Risk quantification, analytics, and management need to be used in a complimentary fashion. Risk quantification and analytics without strong risk management techniques including established governance procedures will fail. Managers need to use the analytics to guide their decision making process. Conversely risk management without strong risk quantification and analytics will also fail.

OPERATIONAL VALUE AT RISK

The concept of value at risk (VaR) was developed by J. P. Morgan in the 1990s as an overall market risk measure. VaR measures the maximum estimated loss in the market value of a portfolio over a specified time horizon with a specified confidence level. This methodology has been adopted for use to quantify operational risk under the advanced measurement approach (AMA) of Basel II, which requires losses over a one-year time horizon at a 99 percent confidence level.

To comply with the AMA method of operational risk capital calculation, banks are spending a great deal of time on working the problems out of their VaR models. This involves developing a separate model for each business line and event type and addressing the following issues:

- Scaling external data to look like internal loss data. Although internal data are the most relevant to the bank, it is generally insufficient to do capital modeling. As a result, Basel II requires banks supplement internal data with external data. Since the internal and external data are generated by different distributions, the external data need to be transformed to look like the internal data.
- Dealing with loss size biases in some of the available databases. Some external loss databases only include losses that are publicly available. These databases only collect data in the tail of the aggregate loss distribution.

EXHIBIT 10.4 Aggregate Loss Distribution

- All operational loss data are collected over a specified threshold level, making it difficult to reliably estimate capital model parameters.
- Estimating the appropriate loss frequency distribution.
- Estimating the appropriate loss severity distribution.
- Combining through convolution or Monte Carlo simulation the frequency and severity distributions with or without a correlation structure.
- Estimating the correlation structure of the data. Business processes are associated with most operational risk. Since business processes are dynamic and not static in nature, estimating correlation for the AMA approach will be difficult at best.
- Banks with weak risk controls are more likely to be represented in external databases because they experience more losses. These banks are also more likely to suffer large losses.

Exhibit 10.4 is an illustration of the aggregate loss distribution that results from the convolution of the frequency and severity distributions, which is used in the AMA approach of capital calculation under Basel II.

Operational value at risk (OpVaR) is useful in estimating the operational risk exposure but does not help manage the risk. Other methodologies are necessary to determine the causal links between process drivers and operational losses.

MULTIFACTOR CAUSAL MODELS

Supplemental models to VaR are needed to understand and identify process drivers in order to manage and reduce operational risk. Most factors that influence operational risk are internal company performance measures. These types of models attempt to explain operational losses with control factors as illustrated in the following equation:

$$Y_t = \alpha + \beta_{1_t} X_{1_t} + \ldots + \beta_{n_t} X_{n_t} + \varepsilon_t \tag{10.1}$$

where Y_t represents the operational loss dollar amount in a particular business line and/or event type; the Xs represent the process drivers; and the α and β_is are

the estimated parameters measuring the impact of the process driver on the loss amount.

To illustrate the possible use of such a model, suppose we have collected data over a six-month period linking operational losses to the following process drivers: computer downtime; time of day in one-hour increments; employee training index; transaction volume; and number of counterparties. The estimated model, once validated for the standard statistical assumptions, can be used to identify influential process drivers (key risk indicators [KRIs] and key performance indicators [KPIs]) and allows management to judge the impact on operational losses due to a reduction in computer downtime.

REGIME SWITCHING MODELS

Regime switching models are commonly used in finance to model processes that move from one state to another over time. For example, the Dow Jones Industrial Index can be modeled as a two-state (expansion and contraction) switching model.

There are two basic approaches taken in regime switching models: thresholds or Markov. The threshold models are generally used when the model's state is believed to follow the observed value of a variable in relation to some threshold. For example, two consecutive quarters of negative real gross domestic product (GDP) growth is the official definition of a recession. The number of consecutive months of negative growth could be used as the threshold. The Markov models are used when the variable that determines the model's state is assumed to follow a Markov process.

A typical Markov switching model is illustrated in equation 10.2.

$$\begin{aligned} Y_t &= X_t \underline{b}_1 \quad S_t = 1 \\ &\\ &X_t \underline{b}_2 \quad S_t = 2 \end{aligned} \tag{10.2}$$

where S_t is the state variable which depends on time and is unobservable. The process determining the state of the model is assumed to follow a Markov process. The probability of moving from state i to state j is determined by the Markov chain:

$$P(S_{t+1} = j | S_t = i) = p_{j,i} \tag{10.3}$$

Consider the company that has experienced a great deal of difficulty in application change management. The company has decided to outsource this function to MBI Computers. Exhibit 10.5 is a plot over time of the monthly downtime minutes for the outsourced application. To assess the financial impact of the decision, management can model the operational losses before and after outsourcing the application change management function. This information will be useful to MBI in deciding whether or not to outsource the change management function for other applications.

Operational risks susceptible to regime shifts will have time varying parameters. More interesting and challenging situations arise when the regime shifts are not single

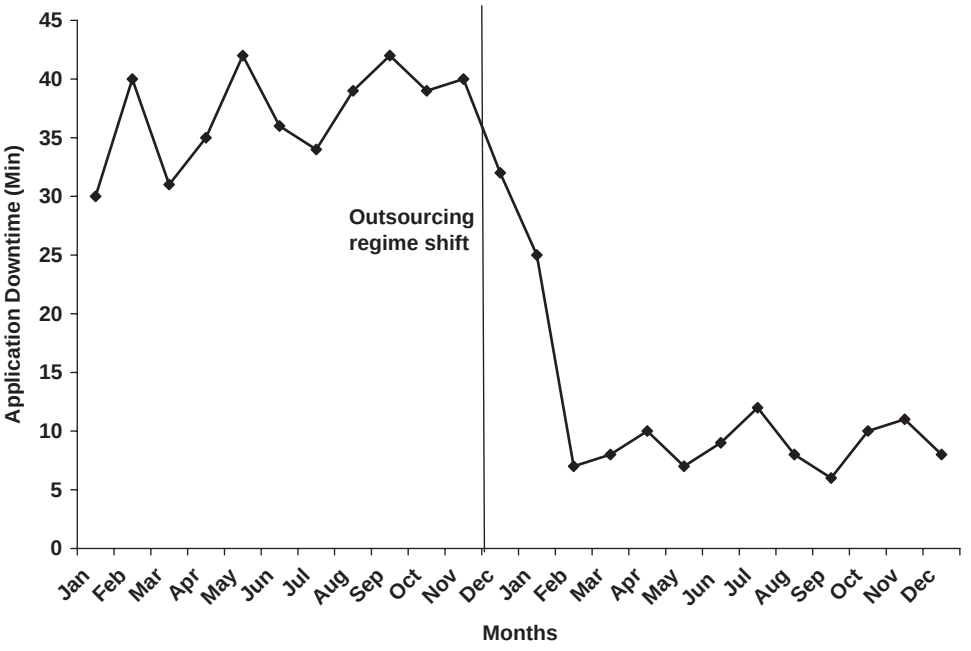


EXHIBIT 10.5 Application Downtime Minutes

deterministic events but are influenced by exogenous events and will occur randomly in the future.

DISCRIMINANT ANALYSIS

Discriminant analysis is a statistical technique for classifying observations into predefined categories. The methodology can be applied to quantitative or ranked qualitative data such as qualitative data from audits and Six Sigma failure analyses. The model parameters are estimated based on a data set for which the categorization of each observation is known. The discriminant function L in equation 10.4 can be written as:

$$L = \mathrm{c} + b_1 x_1 + b_2 x_2 + \ldots + b_n x_n \tag{10.4}$$

Where c is the constant, the b_is are the discriminant coefficients and the xs are the predictor variables. The linear discriminant function can be used to predict the class of a new observation with an unknown categorization. For a situation with two categories, two discriminant functions, L_1 and L_2, are estimated.

$$L_1 = c_1 + b_{1,1} x_1 + b_{1,2} x_2 + \ldots + b_{1,n} x_n \tag{10.5}$$

$$L_2 = c_2 + b_{2,1} x_1 + b_{2,2} x_2 + \ldots + b_{2,n} x_n \tag{10.6}$$

A new observation is categorized into the class for which the discriminant function has the highest value.

BAYESIAN NETWORKS

A Bayesian network (BN) is a graphical model. It reflects the states of the process modeled and integrates the probabilistic relationships between the variables. BNs facilitate modeling cause and effect relationships in complex inference networks. These models can be driven by empirical data and can handle missing data by incorporating expert opinion. Chapter 13, Bayesian Networks for Root Cause Analysis, provides an introduction to this type of model.

PROCESS APPROACH TO OPERATIONAL RISK

Financial firms such as hedge funds have to contend with many risks including market, credit, liquidity, legal, and operational, to name a few. After identifying, assessing, and quantifying risks, it is of utmost importance that there is a process in place to monitor and control or mitigate the residual risk that is present in any of the key processes. Instead of creating a separate process for monitoring and controlling operational risks, it is in the best interest of any organization to pull this process under the organization's overall risk monitoring and controlling strategy. The BCBS recommends that the board of directors ensure that a bank's operational risk management framework is subject to effective and comprehensive internal audit by operationally independent, appropriately trained, and competent staff. The BCBS also states that the internal audit function should not be directly responsible for operational risk management. Organizations should regularly review their risk control strategies to make sure they are effective and help organizations stay within their acceptable risk profile.

Four different process approaches to operational risk will be discussed: business process modeling and simulation; precursor analysis; agent based modeling; and the Six Sigma approach to risk.

BUSINESS PROCESS MODELING AND SIMULATION

Although it may appear that most operational risks are preventable with the implementation of procedures and controls, it is not an easy task to identify and control all risks. An effective method of identifying and ultimately quantifying the operational risk in a company is through business process modeling (BPM) and simulation. For decades, simulation process modeling has been employed in manufacturing and transportation to model physical systems. Recently, this type of process modeling has been applied to business process such as transaction processing and corporate governance processes. The process simulation model will aid the organization in:

- Developing insights into the operations of the business.
- Leverage assets and reduce costs.

- Testing process changes before implementation (change management).
- Experiment with process improvements to reduce cycle times and manage operational risk.
- Conducting stress tests and scenario analysis.

The speed of business today provides only short windows of opportunity. Businesses must bring new products and improvements to market quickly and cannot rely on lengthy, costly, or error-prone projects.

BPM allows a firm to assess the internal structures of the entire organization. It enables the firm to separate processes, systems, and data into distinct layers allowing the firm to monitor them independently. BPM gives the analyst the ability to: determine process performance before and after process changes; perform scenario testing; and stress test the system. For example, a proprietary trading firm interested in moving to stream processing can model the flow of data through the trade generation process. The firm can model the volume of data coming into the system from the data consolidator and can model the time it takes for data to come in the door to the generation of a trade from the trading algorithm.

PRECURSOR ANALYSIS IN OPERATIONAL RISK MANAGEMENT

The nuclear energy and aerospace are two industries that have done extensive research in risk and failure prevention analyses. As a result of catastrophes such as Chernobyl and the space shuttle *Challenger,* a great deal of research has been conducted by both industries to be better able to identify problems in complex systems before they lead to failure. Research has shown the importance of expanding the data set of catastrophic failures with near misses when there are few or zero failures. In 2003, the National Academy of Engineering Program Office initiated the Accident Precursors Project to study accident precursor analysis and management. The Accident Precursors Project defines precursors as "... conditions, events, and sequences that precede and lead up to accidents."[3] They are warning flags of potentially more dangerous situations. Sometimes the warning flags are read correctly and other times they are ignored.

Not all precursors result in catastrophic failures. As illustrated by equation 10.7, failures occur when the precursors are present with additional aggravating factors along with the absence of mitigating factors.

$$\text{Failure} = \text{Precursor} + \text{Aggravating Factor} - \text{Mitigating Factor} \tag{10.7}$$

Near misses are precursors in their own right. They contain very valuable information and should be used as supplemental data to the actual failure events. Unfortunately many near misses are neither recognized nor recorded in some industries.

Many organizations sponsor and support precursor identification. This includes government regulatory agencies such as the Federal Aviation Administration's (FAA's) Aviation Safety Reporting System (ASRS), individual companies in the airline

and nuclear industry, and the medical community's Patient Safety Reporting System (PSRS). Near-miss data are very important and informative for aviation oversight organizations such as the FAA.

A better understanding of operational risk will be gained in the financial services industry when data on risk factors are collected and analyzed. The current focus of the industry is on measuring risk exposure through OpVaR. More important gains will be made when the focus shifts to understanding the drivers behind the risk profile. This will entail analyzing actual loss events and near misses.

AGENT-BASED MODELING

Agent-based modeling (ABM) and simulation has been used since the 1990s in the social sciences to develop new theories and to provide evidence for existing theories. It is a method currently used to study complex systems such as corporations and the stock market. These complex systems cannot be modeled through analytical expressions. ABM models a system as groups of autonomous interacting decision making entities called agents. Each agent is governed by a set of rules that it applies based on the circumstances of the agent. The end result is a distributed decision-making process.

ABM can be used to understand and measure the operational risk of a company through modeling the business processes. A joint project by Icosystem, Bios and Cap Gemini3 applied agent-based modeling to operation risk in the asset-management unit of Société Générale. They modeled employees as interacting agents. Using historical data on losses and errors, the researchers modeled several common mistakes such as confusing local currency with the euro. Through the modeling they were able to discover under what circumstances these common errors led to catastrophic losses. ABM was able to uncover vulnerabilities in the business processes of the bank.[4]

SIX SIGMA APPROACH TO QUALITY AND PROCESS CONTROL: FAILURE MODES AND EFFECTS ANALYSIS

It is not always the case that reliable historical data is available for analysis for any given process within an organization to quantify process failures and the risk induced by these failures. Six Sigma methodology, which has gained a strong foothold in the business community as the most desirable process improvement methodology, relies heavily on data-driven analysis. One of the tools used within Six Sigma to design and implement a robust process is to identify failure modes and establish a risk priority so that corrective actions can be put in place to address and or reduce the risk. This tool is called Failure Modes Effects Analysis (FMEA). FMEAs help in identifying and documenting where in the process the source of the failure impacts the customer (internal or external customer).

FMEA is used to determine failure modes and assess risk posed by the process and thus to the organization as a whole. The first step in the process is to identify key

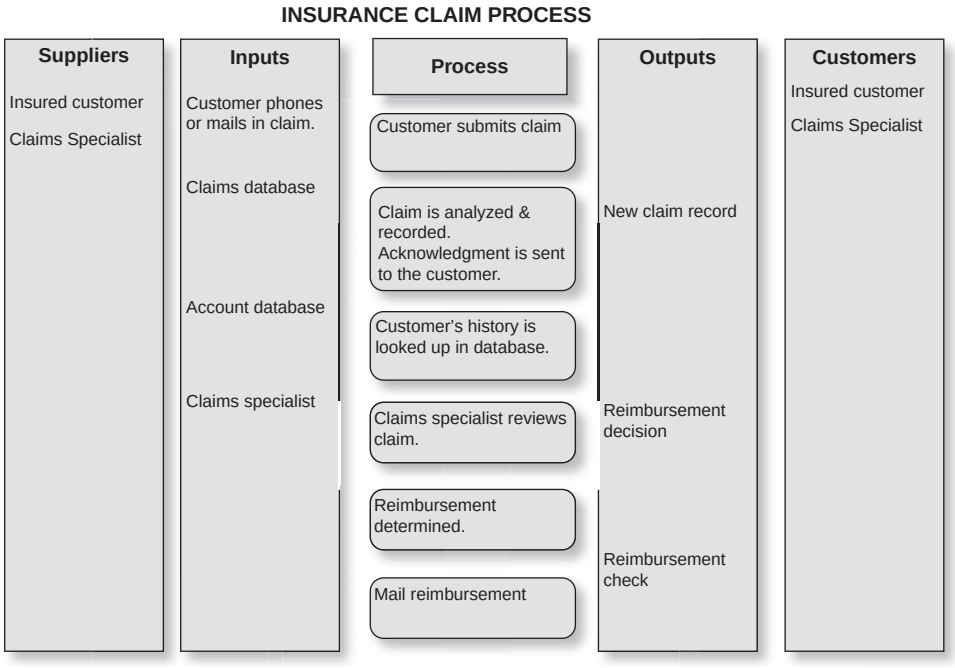


EXHIBIT 10.6 SIPOC Insurance Claim Process

processes within the company or organization. A typical business is comprised of many processes that help run the business and achieve its goals and objectives. Not all these processes are directly related to selling of a product or revenue generating but indirectly contribute to the success of the organization and hence can definitely have an opposite effect as well. Not every process has the same impact, positive or negative, on the business and hence it is important to identify key processes that need to monitored and managed from an operational risk management perspective. The outcome of an FMEA is a risk priority number (RPN). Generally, the higher the RPN, the greater the priority associated with fixing the associated cause of process failure.

The second step is the construction of process maps that graphically illustrate the business process under study, including the interrelationships and dependencies with other processes and departments. This should include conducting, for each subprocess, a SIPOC[5] analysis, which provides a high-level summary of the process. Exhibit 10.6 illustrates a SIPOC diagram for an insurance claims process.

SIPOC diagrams five key elements:

1. *Suppliers*. Roles or people that produce the inputs to the process.
2. *Inputs*. Key process information available before beginning an activity.
3. *Process*. High-level process activities that transforms inputs into outputs.
4. *Outputs*. Key process deliverables.
5. *Customers*. Internal or external users of the process outputs.

The third step is the identification of all potential failure modes in the process and to determine the impact of the failure on the business. The output of this step is the RPN, which is determined by taking into account:

- *Severity*. Each failure is evaluated in terms of the worst possible result of a failure.
- *Likelihood of occurrence*. Each failure is categorized by its likelihood of occurrence. A low number indicates the failure is not very likely, and a high number indicates the failure is very likely.
- *Detectibility*. Each failure is categorized by the likelihood of discovering the failure before the customer is affected. A low rating indicates it is very likely the failure will be discovered early and a high number indicates there is a high likelihood the failure will not be discovered.

The RPN is the product of the severity, occurrence, and detectability category ratings.

The fourth step is to identify corrective actions to eliminate failure or to control risk. As a general rule, a higher priority is assigned to fixing potential process failures with a high RPN.

CONCLUSION

Although operational risk is an issue for all companies, the banking and insurance industries have a focused interest in measuring their operational risk exposure due to the capital requirements of Basel II and Solvency II, respectively. This marks the beginning and not the end of their effort to model and understand their operational risks. Subsequent to their capital modeling effort, attention will turn to understanding the drivers of operational risk so they can effectively manage and change their risk profile.

Developing an understanding of the drivers of a firm's risk profile will entail data collection, data cleansing, and statistical analysis of the drivers. Although risk is sometimes analyzed and measured in silos, it certainly doesn't occur in silos. Risk is multidimensional and needs to be analyzed in this manner. The current subprime credit situation is a perfect example of the multidimensionality of risk. There is no one single factor that caused the current problems—not mark-to-market valuation; not Freddie Mac and Fannie Mae, and not credit default swaps. In-depth analyses are required to understand the interdependencies and to determine the root causes of loss events.

BIBLIOGRAPHY

Bonabeau, Eric. "Agent-Based Modeling: Methods and Techniques for Simulating Human Systems." www.pnas.org/cgi/doi/10.1073, 2002.

Chorafas, Dimitris N. *Risk Management Technology in Financial Services*. Oxford, UK: Elsevier, 2007.

Da Costa Lewis, Nigel, *Operational Risk with Excel and VBA*. Hoboken, NJ: John Wiley & Sons, 2004.
Gallati, Reto. *Risk Management and Capital Adequacy*. New York: McGraw Hill, 2003.
Laguna, Manuel, and Johan Markland. *Business Process Modeling, Simulation, and Design*, Upper Saddle River, NJ: Pearson Prentice Hall, 2005.

NOTES

1. Basel Committee on Banking Supervision, "Sound Practices for the Management and Supervision of Operational Risk," February 2003.
2. Duncan Wood, "Easy Does It," *OpRisk and Compliance* 9(10), 2008.
3. James R. Phimister, Vicki M. Bier, and Howard C. Kunreuthers, *Accident Precursor Analysis and Management: Reducing Technological Risk through Diligence* (Washington, DC: National Academies Press, 2004).
4. Eric Bonabeau, "Predicting the Unpredictable," *Harvard Business Review* 80(3), 2002.
5. SIPOC is a flowcharting method used to illustrate the linkages between suppliers, inputs, process activities, outputs, and customers.

CHAPTER 11

Statistical Process Control Integrated with Engineering Process Control

Deborah Cernauskas, Ph.D., and Bruce Rawlings

INTRODUCTION

Process control is a discipline that deals with monitoring, adjusting, and controlling the output of a process through the use of various methods, procedures, and algorithms. While control systems can be found throughout history, a formal discipline was not developed until the late 1800s or early 1900s, and it is only recently that applications in business process controls have taken root. Although classical control theory is more commonly associated with manufacturing, the methods, procedures, and algorithms can be transformed and applied to improving the quality of and reducing the losses associated with the business processes of any firm.

Two main branches of process control have developed in different industries over time. Statistical Process Control (SPC) originated in the parts manufacturing industry, and Engineering Process Control (EPC) in the process industry. SPC has been employed extensively to monitor and control processes through the use of control charts and focuses on eliminating the root cause of variability. SPC tries to improve the process over the long run. EPC, however, focuses on controlling the drivers of the process to ensure quality. EPC is a short-term approach that attempts to minimize process variation by transferring the variation into another variable.

An illustration will aid in understanding the difference between SPC and EPC. Suppose you are an equities portfolio manager. Your investment goal is to construct the portfolio to achieve a certain level of return while controlling the level of risk. SPC considers the process in control if the Sharpe ratio, measuring the risk-return trade-off, does not differ significantly from the desired Sharpe ratio (setpoint). To go from the desired to the actual Sharpe ratio, the portfolio manager must make adjustments to the process variables (e.g., equities within the portfolio and the level of investment in each equity).

EPC is focused on the process variables that are affected by externalities (e.g., currency and interest rates) that cannot be controlled by the portfolio manager. EPC

makes adjustments to these process variables to maintain the desired risk-return trade-off.

Although SPC and EPC share a common goal—quality assurance—they approach the issue from alternative directions. SPC stresses the oversight of processes and fault recognition with minimal process adjustments. SPC is most effective when the process outputs are independent and identically distributed (IID) and the quality goal is to find departures from this assumption. Fine-tuning the process when it is statistically in control will only result in increasing the process variation instead of a reduction. EPC advocates parameter fine-tuning to keep the process from drifting too far away from the target performance measure. Controlling individual parameter values through EPC will not guarantee that the process will not drift out of control. Vander Wiel et al. found an integrated approach to quality using both SPC and EPC leads to process optimization and process improvement.[1]

Many financial processes are inherently hard to control for several reasons. First, there is generally a time delay between input variable changes and when an effect is observed in the system output. For example, consider a bank that issues consumer loans. The bank relies on input data concerning the creditworthiness of the loan applicant. If the bank makes a mistake in issuing credit to a noncreditworthy applicant, the bank does not know immediately. Second, many of the processes are affected by externalities such as market prices that cannot be controlled.

The remainder of this chapter provides an overview of control schemes and a discussion of common SPC and EPC methods. Additionally, it provides an example of an integrated SPC/EPC control system in a trading environment.

CONTROL SCHEMES

The processes of interest in this chapter are information processes such as those used to process trades at a hedge fund, payroll, financial plans and budgets, and inventory. In today's business environment, even small firms use computer systems to drive these processes. Information processing runs virtually everything, including governments, manufacturing plants, medical services, hedge funds, financial markets, and transportation systems.

The information processing involved in running a business, for example, is in reality a complex network that is influenced and driven by many factors. The network surrounding and embedding a process may include other internal processes; external information arriving through the mail, e-mail, or downloads; and internal information arriving through e-mail and downloads. Ensuring the integrity and accuracy of an information process is not an easy task but can be improved through the application of control schemes such as SPC and EPC, which are naturally complements to each other. Through the application of control charts, SPC identifies departures from the presumed process model. However, EPC is intended to work within the process model by adjusting the process drivers for expected types of process disturbances.

In many control systems, a large number of parameters are simultaneously monitored. Traditional SPC techniques assume the parameters are independent and are geared toward identifying abrupt process changes. Alternatively, EPC control systems are less effective on IID processes and more effective on trending processes. EPC

and SPC techniques can nicely complement each other and provide a more robust quality system.

Suppose $\underset{\sim}{X}_1, \underset{\sim}{X}_2, \ldots \underset{\sim}{X}_t, \ldots$ is a sequence of observations related to an information process. Each $\underset{\sim}{X}_i$ represents a vector of measures taken on the process at time i. A typical control scheme for a process is a set of criteria that enables one to judge if the process is in control; that is, are the values within an acceptable level of variation compared to a target value? At some point the process will signal an out-of-control state emanating from a change in one or more underlying parameters or from randomness in the data.

Different control systems use different methods to adjust an out-of-control process. Two common EPC control systems are feedforward and feedback systems. Feedforward systems are designed and used to thwart errors from entering or disturbing a process. Alternatively, feedback systems are used to correct errors that have already occurred and are detected within the process. Consider the analogy of a residential burglar alarm system. A feedforward system would turn on outside lights when someone approaches the residence too closely. A feedback system does not take any action until the burglar is already in the house and then someone dials 911.

STATISTICAL PROCESS CONTROL

The principles and methodologies of SPC are not industry dependent and rely on simple sample statistics (mean, range, and standard deviation) to analyze data. Control, cumulative sum control (cusum), and exponential weighted moving average charts are the vehicles used to monitor these statistics.

Control chart theory is the exact opposite of statistical process modeling. The statistical approach fits the model to the process, while the control chart approach fits the process to the model. Control chart theory assumes the empirical data are IID. Factors that cause the process to act differently are deemed special-cause variation and require immediate attention and elimination.

SPC attempts to answer two main questions: (1) is the process under control?, and (2) does the process meet the intended specifications?

SPC Tools

Exhibit 11.1 illustrates the common tools used in SPC. Most of these tools have been around for decades. They have the advantage of being easy to implement and interpret.

Data collection and presentation tools give the analyst tools to make judgments on data quality and gain insights into the data. Exploratory data analysis (EDA) is a term developed by Tukey to describe the process of looking at numbers and graphs to find patterns and structures in data.

Problem-solving tools provide an integrated picture of an information process identifying departmental interdependencies and data dependencies (process mapping and flowcharting), and the means of determining and graphing the set of possible root causes of a problem (cause-and-effect diagram and Pareto charts).

Descriptive statistics are summarization tools that help the analyst determine distributional properties of empirical distributions. There are three main

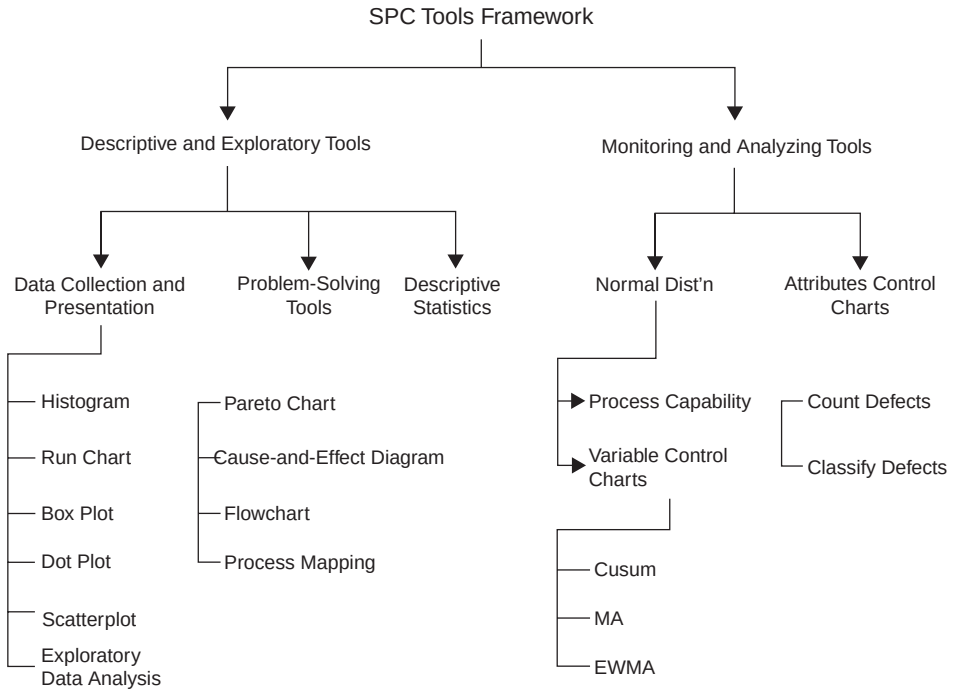


EXHIBIT 11.1 SPC Tools Framework

characteristics of interest: measure of central tendency (mean, median, and mode), dispersion (range, variance, and standard deviation), and shape (skewness and kurtosis).

Control Charts

Control charts are the most commonly used SPC technique and are very easy to implement. Control charts provide a historical record of the performance of a process, which, when combined with business process modeling, can be used to understand the impact of proposed process improvements.

Control charts for individual process measures are used to determine if a special cause variation caused the central tendency of the process measure to change or drift over time—its lower control limits (LCL) and upper control limits (UCL).

$$LCL = \overline{X} - 2.66 \times \overline{R} \tag{11.1}$$

$$UCL = \overline{X} + 2.66 \times \overline{R} \tag{11.2}$$

$$R = \text{Subgroup max} - \text{Subgroup min} \tag{11.3}$$

where $\overline{X}$ is the mean of the observed process measures; $\overline{R}$ is the average range value; and 2.66 is the value used for individual measurement plot with range subgroups of two observations. See Exhibit 11.2 for an example control chart for an individual process measure.

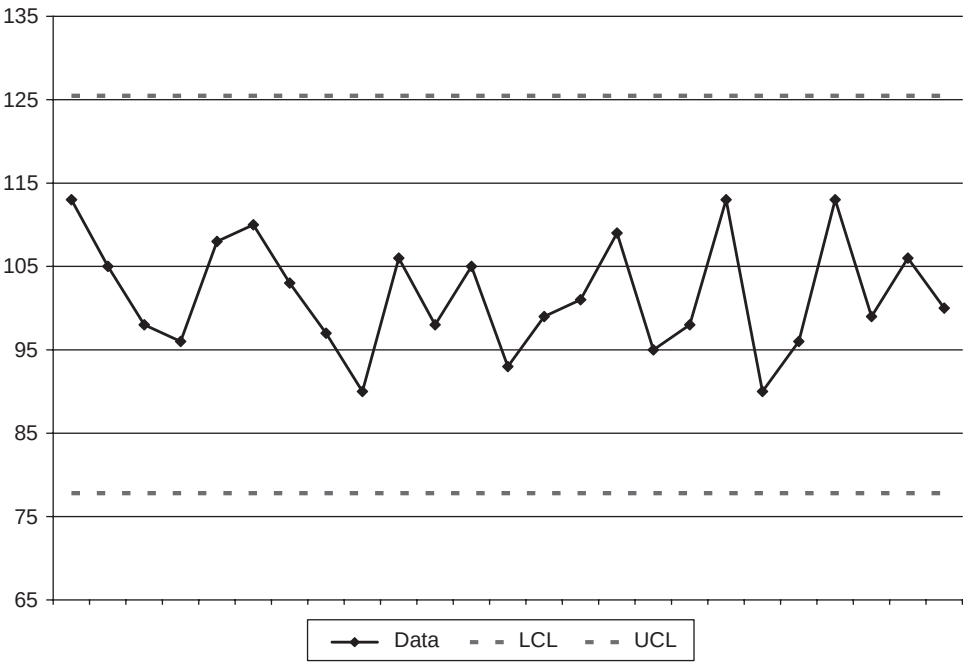


EXHIBIT 11.2 Example Control Chart for an Individual Process Measure

ENGINEERING PROCESS CONTROL SYSTEMS

EPC and Automatic Process Control (APC) are terms used to describe control systems, which implement adjustments to process drivers. Three common types of EPC systems include:

1. On-off control
2. Open-loop control
3. Closed-loop control

Each of these systems will be discussed briefly.

On-Off Control

On-off control has the longest history of use and is the crudest form of control system. There are at least four types of variables in EPC systems: setpoint (SP); process variable (PV); manipulated variable (MV); and disturbances. The setpoint is the desired value of the system output. The process variable describes what we are trying to control. The manipulated variable(s) is a variable in the process that can be changed in order to keep the process functioning within the desired range. The disturbances are inputs to the system that cannot be controlled.

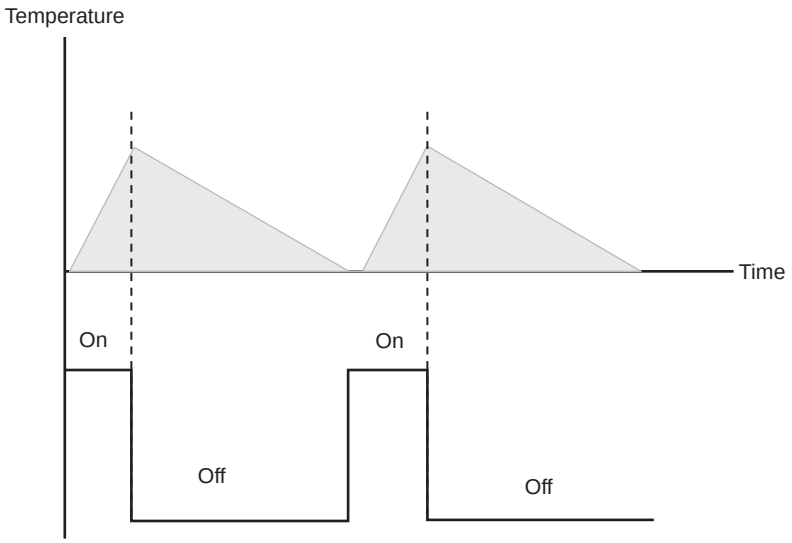


EXHIBIT 11.3 Residential Home Heating System Example
Source: Adapted from Wolfgang Altman, *Process Control for Engineers and Technicians*. Amsterdam: Elsevier, 2005.

Mechanical On-Off Control Example On-off controllers are commonly found in residential appliances such as a thermostat (see Exhibit 11.3). The setpoint is the desired room temperature; the process variable is the actual room temperature; the manipulated variable is the fuel flow into the furnace; and a disturbance is the outdoor temperature.

On-off controllers are easy to execute and economical. This method is feasible only when large variations in PV are acceptable.

Finance On-Off Control Example Every portfolio manager is familiar with Markowitz's mean variance (MV) optimization for asset allocation. Although taught universally in business schools, the methodology is rarely implemented in practice because it has a severe limitation in its sensitivity to small changes in the inputs.

Michaud's Resampled Efficiency^{TM} (RE) technique is a method of dealing with parameter estimation error in computing the MV optimal portfolio when rebalancing positions.[2]

According to MV theory, investing in any portfolio below the efficient frontier is suboptimal. An investor holding portfolio C in Exhibit 11.4 has the same level of risk as portfolio B but a lower return. Similarly, portfolio C has the same return as portfolio A but a higher level of risk. Any portfolio on the efficient frontier will provide the investor with the optimal risk-return trade-off. The theory is very logical but hits a speed bump during implementation. Errors in estimating the expected return and standard deviation of the various portfolios are not considered. As illustrated in Exhibit 11.5, the portfolios within the ellipse are statistically as efficient as those on the efficient frontier due to estimation error.

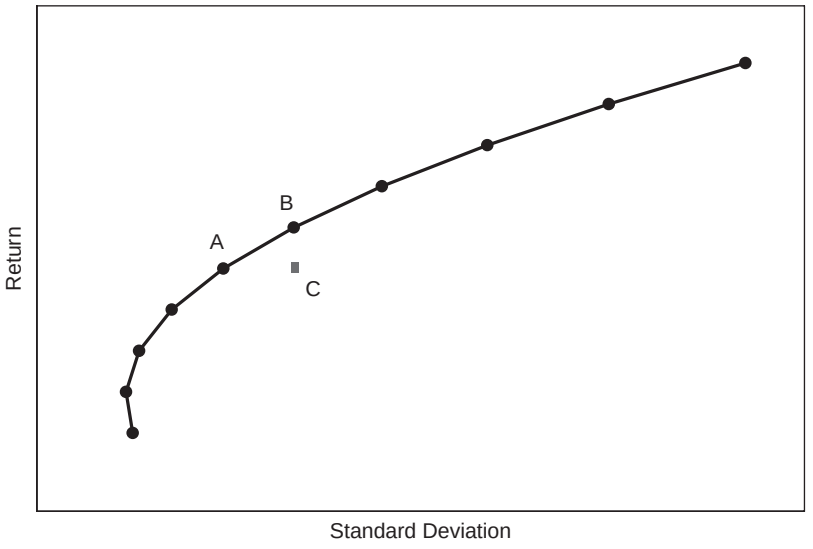


EXHIBIT 11.4 Example Efficient Frontier

An interesting question raised by Jobson and Korkie is: "Does the efficient frontier have a variance?"[3] Since the MV efficient frontier is based on statistically estimated parameters, the answer has to be yes. Both Michaud, Jobson, and Korkie have developed a statistical equivalence region through resampling techniques. This concept can be used as the basis for an EPC quality control system in which the portfolio rebalancing occurs only when the risk-return trade-off falls outside of the statistical equivalence region (see Exhibit 11.6).

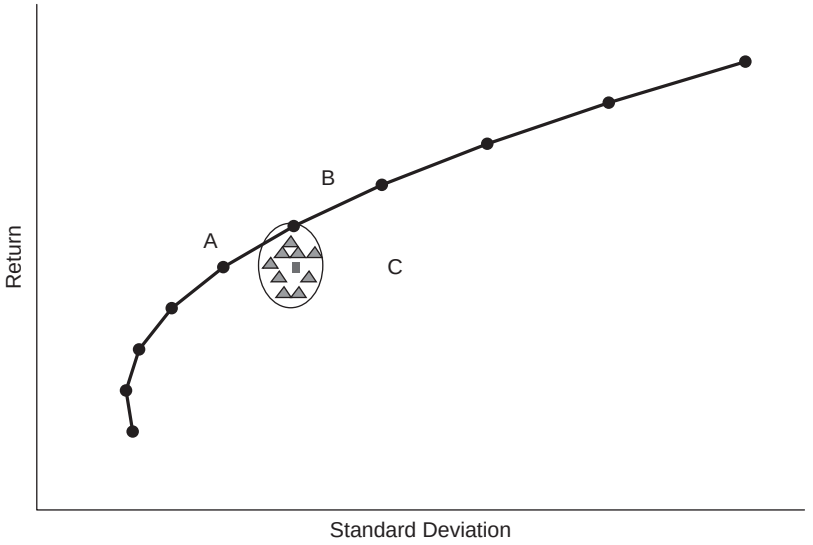


EXHIBIT 11.5 Statistically Efficient Portfolios

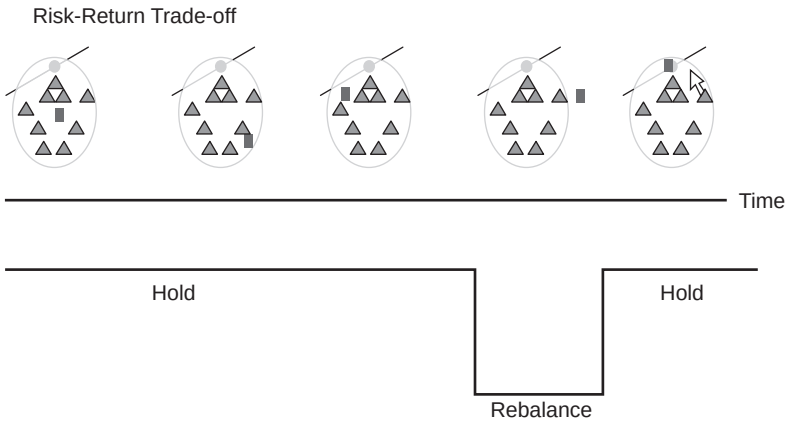


EXHIBIT 11.6 Portfolio On-Off Control

Open-Loop Control

The most common type of open-loop control system is feedforward control. This technique bases control on the state of the disturbances without regard to the state of the system output. The input variables are adjusted to compensate for the impact of the process disturbances. This type of system results in fast corrections to the system but requires a good understanding of the effects of disturbances on the system.

Closed-Loop Control

Control action in a closed-loop control system (aka feedback system) is determined by the state of the PV. These systems are designed to maintain the system at the setpoint value. The controller's corrective action is determined by the magnitude of the difference between PV and SP.

Exhibit 11.7 is a block diagram illustrating a single-loop feedback control system. The setpoint is an input value that represents the desired process output. The setpoint in the portfolio example is the targeted risk-return trade-off; a process variable is the Sharpe ratio, which measures risk and return; manipulated variables

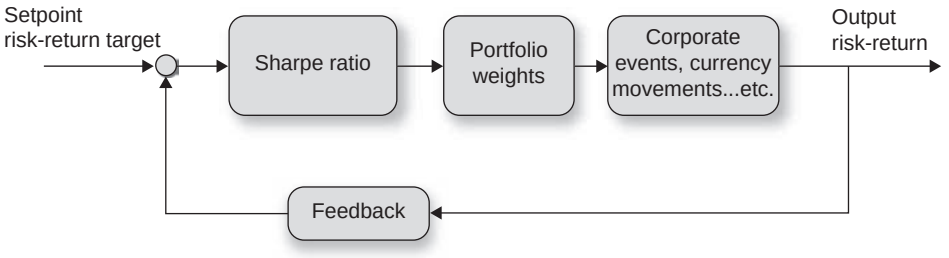


EXHIBIT 11.7 Single-Loop Feedback System

EXHIBIT 11.8 Control Modes

Control Mode	Equation*
Proportional	$P_t = P_0 + K_c e(t)$
Integral	$P_t = P_0 + \frac{K}{\tau}\int_0^{\tau} e(t)dt$
Derivative	$P_t = P_0 + \tau_D \frac{de(t)}{dt}$

*P_0 describes the process output when the disturbance variable, $e(t)$, is zero. Kc is the controller gain and describes the size of the process correction based on the size of the process error, $P_t - SP_t$.

include the portfolio weights; and disturbance variables include the corporate events such as stock splits and mergers, currency fluctuations, earnings growth, and so on.

In general, there are three control modes available in feedback systems. The modes are proportional (P), integral (I), and derivative (D). These modes can be combined or used separately in a feedback system (see Exhibit 11.8).

EPC Summary

A disadvantage of the feedback system is that no action is taken until a large deviation in one of the controlled variables occurs. A measurable and significant error is necessary to prompt any action. There are three common criticisms of EPC:

1. Results in overcompensation for process disturbances.
2. Compensates for disturbances instead of eliminating them.
3. Obscures process information that may be used for quality improvements.

EPC results in a more competent process but in the long run does nothing to improve the underlying process.

FINANCE EXAMPLE

Consider an equity market neutral hedge fund that buys and sells stocks with the goal of neutralizing exposure to the stock market by neutralizing beta. The fund seeks to generate returns by exploiting stock market inefficiencies. The strategy tries to generate positive returns in both bull and bear markets and uses a proprietary trading model for selecting trades. The portfolio positions will be adjusted as the market changes to keep beta close to zero (see Exhibit 11.9).

In this example, the portfolio is comprised of two stocks, BNI and IBM. The portfolio is formed on January 2, 1981. The initial betas for the individual stocks are estimated from the following regression equations:

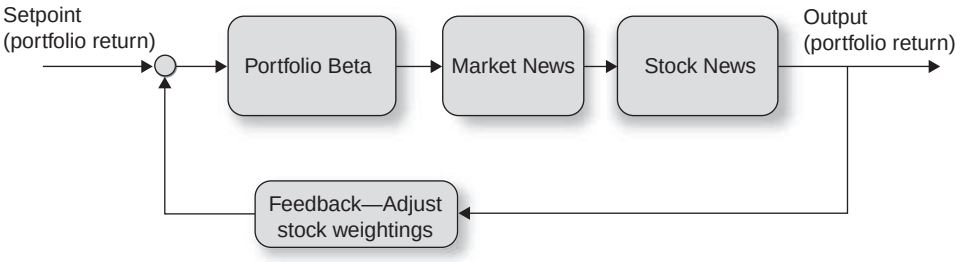


EXHIBIT 11.9 Beta Neutral Portfolio Strategy

$$\begin{aligned} r_{BNI,t} - r_{f,t} &= \alpha_{BNI} + \beta_{BNI}(r_{mkt,t} - r_{f,t}) + \varepsilon_{BNI,t} \\ r_{IBM,t} - r_{f,t} &= \alpha_{IBM} + \beta_{IBM}(r_{mkt,t} - r_{f,t}) + \varepsilon_{IBM,t} \end{aligned} \tag{11.4}$$

The portfolio beta is estimated as:

$$\hat{\beta}_{Port,t} = w_{BNI,t} \times \hat{\beta}_{BNI,t} + w_{IBM,t} \times \hat{\beta}_{IBM,t} \tag{11.5}$$

Exhibit 11.10 clearly shows the portfolio beta is very volatile over time and needs to be rebalanced to keep the portfolio market neutral. The typical Shewhart control chart as illustrated in Exhibit 11.11 is commonly used in SPC and uses control limits based on $\bar{x} \pm 3s$. Shewhart charts are not powerful in detecting small changes in the

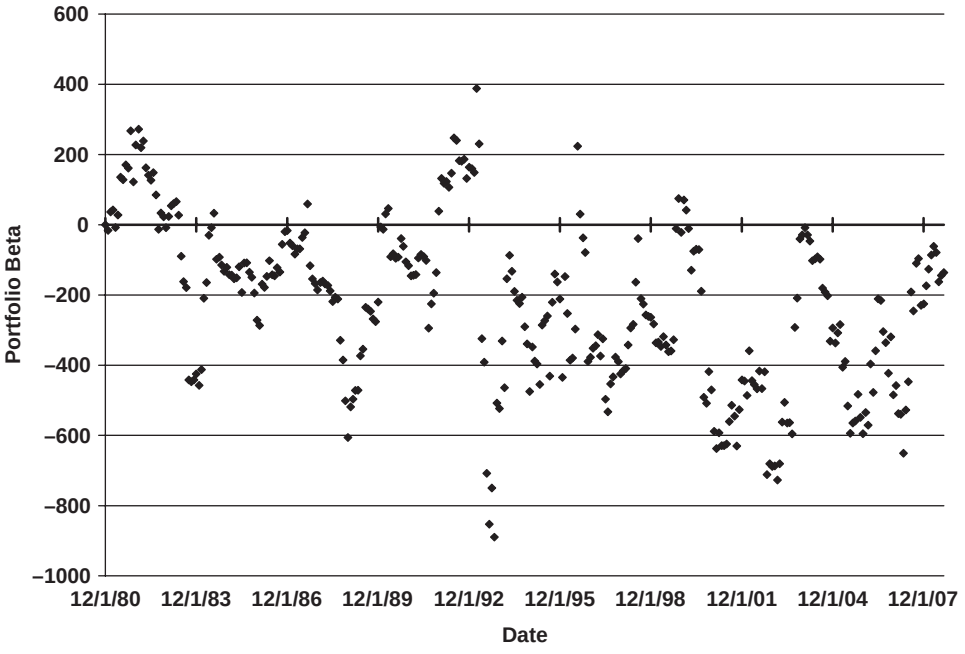


EXHIBIT 11.10 Portfolio Beta with No Rebalancing Over Time

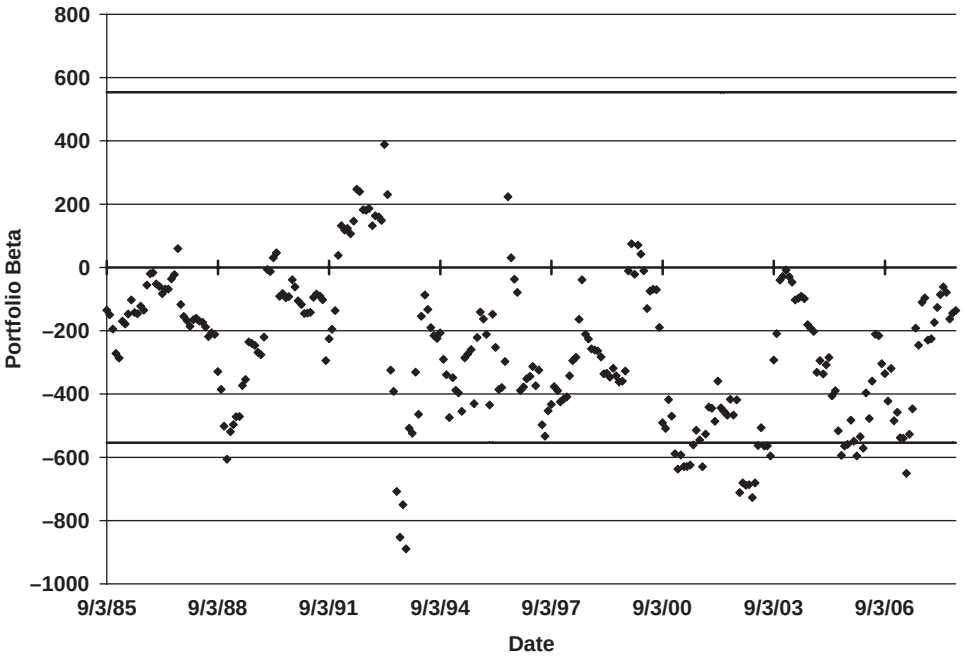


EXHIBIT 11.11 Shewhart Control Chart

process. The mean and standard deviation used to construct the upper and lower control limits were based on the monthly data from 12/1/1980 through 12/2/1985. The standard Shewhart control chart is not very sensitive to changes in the data. Exhibit 11.6 clearly shows the portfolio beta trending negative. It also takes a long period of time before the Shewhart chart picks up the trend.

The cusum chart in Exhibit 11.12 shows a clear pattern in the data—the portfolio beta is becoming more negative over time.

An SPC system will focus on monitoring the portfolio return or, alternatively, the portfolio beta to identify when the system is out of control. An EPC system will focus on monitoring and controlling the underlying drivers. In this example, an EPC system will monitor and place controls on the individual stock betas.

The portfolio manager's goal is to keep the portfolio beta as close to zero as possible without generating excessive and unnecessary trades. Using an SPC system, monitoring and controls are placed on the portfolio beta. The portfolio weights are not adjusted when the calculated portfolio beta falls within three standard deviations of zero (equation 3). Alternatively, the portfolio weights are adjusted when the portfolio beta falls outside of a three standard deviation range.

$$\hat{\beta}_{port,t} = \hat{\beta}_{port,t-1} \text{ when } -3^{*}std_{\hat{\beta}_{port}} < \hat{\beta}_{port} < 3^{*}std_{\hat{\beta}_{port}} \tag{11.6}$$

Implementing this process yields the portfolio betas as illustrated in Exhibit 11.13.

Alternatively, monitoring and control can be applied to the process drivers, in this case the individual stock betas, and the portfolio beta. Exhibit 11.14 illustrates

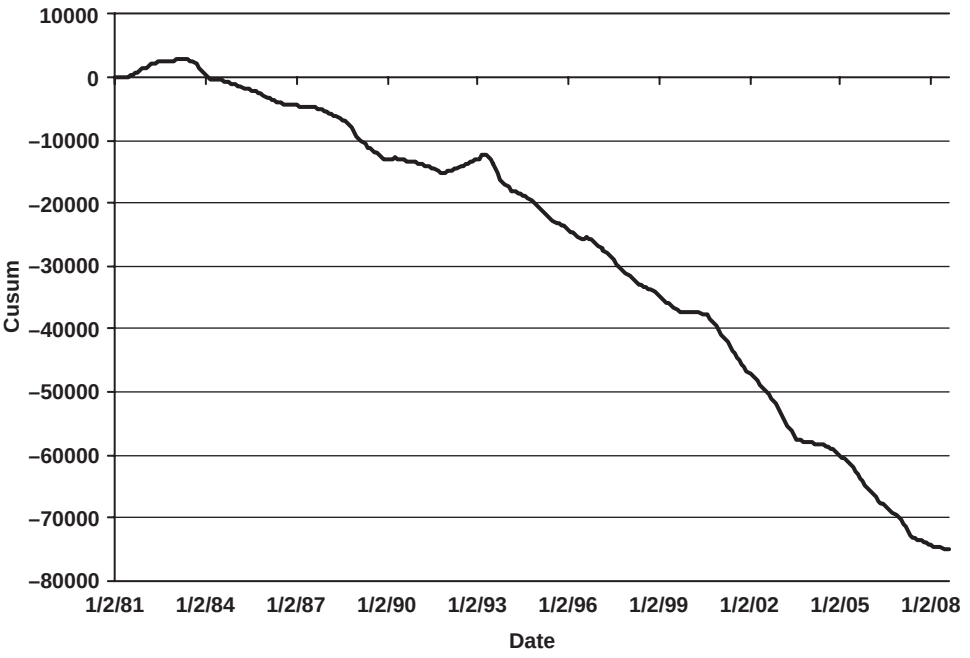


EXHIBIT 11.12 Cusum of the Monthly Portfolio Beta

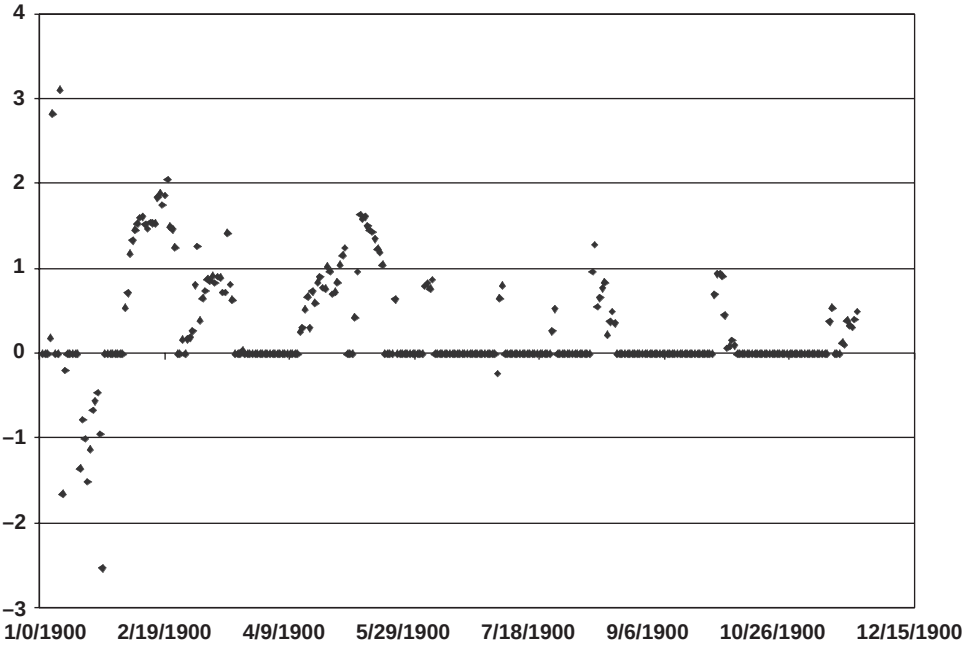


EXHIBIT 11.13 Portfolio Betas with SPC Monitoring and Control

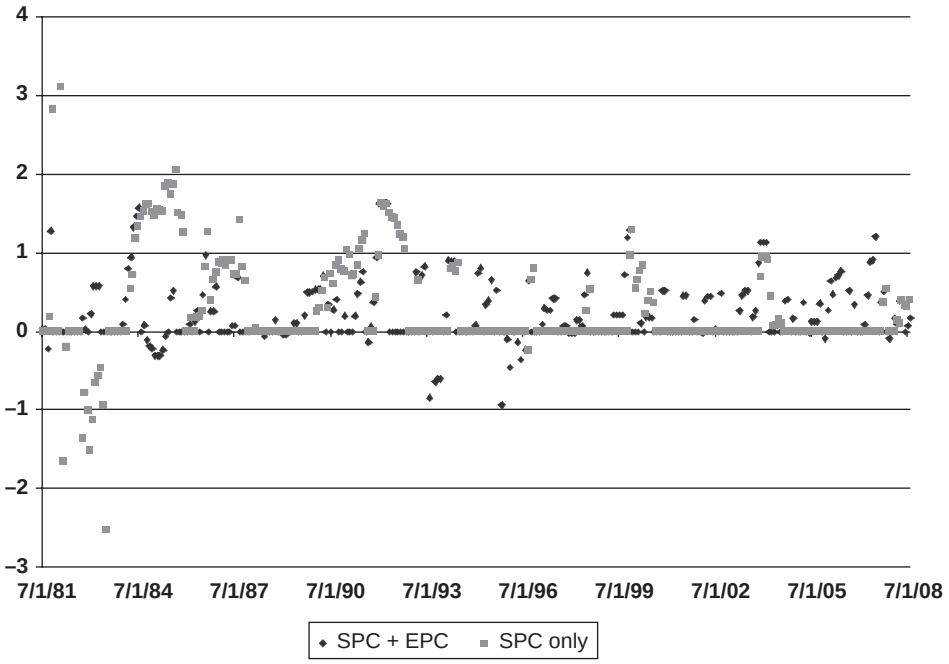


EXHIBIT 11.14 Portfolio Betas Using SPC Only and Combined SPC/EPC Monitoring and Control

a combined SPC/EPC monitoring and control system. The individual stock betas are not adjusted when

$$\beta_{t-1}^{stock} - 3\sigma_{\beta,t-1}^{stock} < \beta_t^{stock} < \beta_{t-1}^{stock} + 3\sigma_{\beta,t-1}^{stock} \tag{11.7}$$

Otherwise, the betas are adjusted according to:

$$\tilde{\beta}_t^{stock} = \hat{\beta}_{t-1}^{stock} + \gamma\left(\hat{\beta}_t^{stock} - \hat{\beta}_{t-1}^{stock}\right) \tag{11.8}$$

which constitutes the monitors and control for the process drivers (EPC).

The preliminary portfolio beta is then:

$$\hat{\beta}_{portfolio,t} = \omega_{t-1}^{BNI}\tilde{\beta}_t^{BNI} + \omega_{t-1}^{IBM}\tilde{\beta}_t^{IBM} \tag{11.9}$$

Additionally, the stock weightings are adjusted when:

$$|\beta_{port}| > \lambda \text{ and } \frac{\omega_{t-1}}{\omega_t} > \tau \tag{11.10}$$

Applying both EPC and SPC monitoring and control results in a more stable process. The standard deviation in the portfolio betas was reduced by 63 percent by adding SPC controls to a process only using EPC controls and by adding EPC controls to a process only using SPC controls (see Exhibit 11.15).

EXHIBIT 11.15 Process Control Comparison

Statistic	EPC only	SPC only	EPC/SPC
Mean	−0.0704	0.2714	0.1971
Maximum	1.7223	3.1091	1.6347
Minimum	−6.6766	−2.5293	−0.9283
Standard Deviation	0.6061	0.6094	0.3819

CONCLUSION

The integration of Statistical and Engineering Process Control exploits the strengths of both systems. While either control methodology will lead to better process quality, the integration of the two approaches will simultaneously optimize the process through process driver adjustments (EPC) and provide long run process improvement through the elimination of the root causes of variability indicated by SPC monitoring.

Six Sigma methodologies are currently being implemented at many large global banks. Not much is seen in the trade press about the quality programs because they are viewed as a means of competitive advantage.

The intent of this chapter was to spark your imagination as to possible ways to apply process control. The implementation of quality techniques are proven to reduce operational losses, reduce rework, increase profitability, and perhaps to reduce the occurrence of rogue trader scandals.

BIBLIOGRAPHY

Box, George, and Tim Kramer. "Statistical Process Monitoring and Feedback Adjustment—A Discussion." *Technometrics* 34(3): 251–267.

Lowry, Cynthia A., and William H. Woodall. "A Multivariate Exponentially Weighted Moving Average Control Chart." *Technometrics* 34(1): 46–53.

Palm, A. C. "SPC versus Automatic Process Control," *Transactions of the 44th Annual Quality Congress* (1990): 694–699.

Rao, Ming, and Haiming Qiu. *Process Control Engineering*. Amsterdam: Gordon and Breach Science Publishers, 1993.

NOTES

1. Scott A. Vander Wiel, William T. Tucker, Frederick W. Faltin, and Neclip Doganaksoy, 1992, "Algorithmic Statistical Process Control: Concepts and an Application," *Technometrics* 34(3) (1992): 286–297.
2. Richard Michaud, *Efficient Asset Management* (Cambridge, MA: Harvard Business School Press, 1998).
3. J. D. Jobson and Bob Korkie, "Putting Markowitz Theory to Work," *Journal of Portfolio Management* 7(4) (1981): 70–74.

CHAPTER 12

Business Process Management and Lean Six Sigma: A Next-Generation Technique to Improve Financial Risk Management

Anthony Tarantino, Ph.D.

BACKGROUND

Business process management and *business process modeling* are terms in popular use and based on the use of electronic workflows as tools to improve processes—often defined as greater efficiencies such as lower costs and shorter cycle times. The same processes can be effective in reducing financial risk management and, when coupled with Lean Six Sigma, can be seen as a next-generation best practice. Let's start with some basic definitions.

A business process is a set of coordinated activities and tasks performed by people, equipment, and computers designed to achieve specific objectives of an organization.

Business process management (BPM) is a poor name for a systematic approach to improve business processes. It is a poor name in that the name implies it is a means to simply manage an existing process. BPM is often associated with technology tools to improve activities, which is not always the case. It is now common for software providers to use the term *BPM* to describe electronic workflows that apply a routing for tasks and activities, including automated controls.

Process models are processes of the same nature that are classified together into a model—a process at the model level. Process models may be used to demonstrate how a process could or should be done as opposed as how the process currently works. It describes how a given process will function. Process models are used to achieve three major goals:

1. *Descriptive*. Tracking the current process and suggested improvements in it.
2. *Prescriptive*. Defining rules and guidelines to achieve the desired processes.
3. *Explanatory*. Providing the rationales for the changes in processes.

Business process modeling compares the current or as-is state of a given process with a desired future or going-to state of process. The goal is evaluate and improve the current state.

Business process improvements will typically require information technology improvements, with the exception of some Just-in-Time or Lean manufacturing improvements, which often are achieved with simple visual controls.

BPM governance is described by Andrew Spanyi, author of "More for Less: The Power of Process Management," as follows: "In order to optimize and sustain business process improvements it's essential to overlay some form of governance that creates the right structures, metrics, roles and responsibilities to measure, improve and manage the performance of a firm's end-to-end business processes. This is called BPM Governance."[1] He argues that it is vital to overlay a form of corporate governance that empowers the appropriate organizational framework, and rules with a system of measurements and alerts to manage an organization's end-to-end business processes. Creating a BPM governance framework should be the first step in any BPM development, and before attempting to find the fastest and cheapest way to get from point A to point B. It would include enterprise-wide collaboration across functions and locations that enforce management accountability and compliance to all appropriate laws, regulations, and standards. Therefore, proposed business process models should be reviewed by the chief risk officer (CRO), chief compliance officer (CCO), and internal auditors, before going into production.

Six Sigma is the letter of the Greek alphabet used to represent the standard deviation of any process.

A Six Sigma quality level is said to represent 3.4 defects per million opportunities. Six Sigma began as the use of statistical methods to improve quality, business process efficiencies, and profitably. Today, it is a methodology for continuous improvement in customer satisfaction and profitability that goes far beyond reducing defects to focus on general business process improvement. While it began and has been used extensively in manufacturing, Six Sigma is gaining wide acceptance in nonmanufacturing sectors. The leaders in the financial services industry have eagerly sought out Six Sigma black belts from manufacturing to develop their own continuous improvement programs. Six Sigma is successful because it utilizes trained, certified, and highly focused experts (green, black, and master black belts) who apply proven methodologies and protocols for problem solving that always defines success in a quantifiable manner and within a defined time frame, typically three to six months. While much has been touted about the use of its hard statistical and mathematical tools, it also applies softer problem-solving and facilitation tools that have proven to be very effective as well. The key is knowing what combination of techniques to apply to a given situation. Six Sigma maintains the following:

- Continuous efforts to achieve stable and predictable process results by reducing process variations that are essential to business success.
- Business processes have characteristics that can be measured, analyzed, improved, and controlled.
- Achieving sustained quality improvement requires commitment from the entire organization, particularly from top-level management.

Common misconceptions and the truths about Six Sigma include:

- *Six Sigma is applicable to manufacturing processes only.* It is widely used outside of manufacturing, and especially in leading financial services organizations.
- *Six Sigma projects require extensive training and certification while creating a major bureaucracy.* Green belt training can be as short as one week, and black belts require a few additional weeks, plus credit for a Six Sigma project. Black Belts do not require any bureaucracy and need not be a dedicated resource.
- *Six Sigma projects are not cost-effective.* To the contrary, Six Sigma projects include a return on investment (ROI) analysis and justification and always produce quantifiable metrics to define success. Since they are typically short term and process oriented, they can generate great ROIs.

Lean Six Sigma is growing in acceptance as a method that combines the best of Six Sigma and Lean thinking/manufacturing. *Lean* is a popular term to describe the Just-in-Time (JIT) techniques created by Taiichi Ohno and advocated by Shigeo Shingo for Toyota in the 1950s and 1960s. Lean strives to eliminate wastes of all kinds—in time, materials, processes, systems, and so on. Six Sigma strives to answer the voice of customers, both internal and external, with a systematic approach to problem solving by certified professionals. Both realize that process improvements are continuous and require ongoing monitoring, analysis, and action. Combining the two provides a powerful best practice for process and customer services improvement and in the hands of focused experts.

Workflows or *electronic workflows* describe a sequence of operations, tasks, or work by one or more people or equipment. Electronic workflows typically include approvals for critical decisions. Many times, electronic workflows are combined with electronic forms creating automated controls and a complete audit trail—very desirable to regulators, auditors, and risk managers. Approval workflows also include time-outs and alternate routings so that delays in a process can be alerted and addressed. The nature of electronic workflows and electronic forms promote standardization over manual processes and controls. They also provide a very visible means to identify redundant or ineffective processes.

HISTORICAL PERSPECTIVE

It is easy to forget that we stand on the shoulders of giants in continuous process improvement. Continuous process improvement and cycle time reductions are as old as man. Many involved breakthrough technologies, but many others were made by the careful evaluation of a process and actually reduced the use of technology.

In the early twentieth century, Frank Gilbreth and his wife, Lillian Gilbreth, pioneered time and motion studies that stressed continuous improvement. Lillian is considered the first lady of engineering, who developed modern industrial engineering. The Gilbreths are household names because of biographical books and the popular movies of the 1950s. *Cheaper by the Dozen* is about their raising 12 children. Their goal was to make the workplace safer and more humane, unlike Fredrick Winslow Taylor, a contemporary of the Gilbreth's, who sought only to squeeze out more work from employees.

Henry Ford revolutionized manufacturing in the 1920s with an obsession with efficiency improvements. Taiichi Ohno, the father of the Toyota Production System, is arguably the greatest contributor to our modern programs to eliminate waste through JIT, Lean manufacturing, and Lean thinking. Ironically, Ohno was inspired by stories of a modern marvel in the United States—the supermarket. Ohno later took Toyota to the next generation by studying possibly the greatest balance of customer service and efficiency: 7 Eleven. JIT and Lean are critical in that they often replaced infective technology solutions with simple visual and process controls. This is not to say that Lean does not rely on sophisticated technologies. To the contrary, it usually involves leading-edge technologies in planning, point-of-sale, radio frequency identification (RFID), statistical process control, and so on.

Six Sigma was developed by Bill Smith of Motorola in 1986 and heavily influenced by decades of earlier quality improvement programs such as Total Quality Management (TQM), Quality Circles, and Zero Defects, as well as the writings of Shewhart (statistical quality control), Deming (14 points and 7 deadly sins) Juran (Pareto Principle), Ishikawa (cause-and-effect diagram), Taguchi (design of experiments), and others. Ironically, many of these approaches became very popular as a perceived cure-all, but fell into disfavor when results were not sustainable.

BPM IN FINANCIAL SERVICES—FUNCTIONALITY TO LOOK FOR

A good BPM solution in financial services should include the following elements and functionality. Many of these elements are common to most any industry sector, but are critical in such a highly regulated, litigious industry in which proper risk/opportunity management is the key business driver.

Business Process Modeling

- Overlay BPM governance frameworks and templates over proposed/planned workflows.
- Play "what if?" with proposed process flow changes.

Business Process Management

- Enforce business rules that provide preventive automated controls, which are the most desired by auditors, reduce audit costs, and improve risk management.
- Provide clear and end-to-end audit trails that satisfy internal and external auditors.
- Automate complex workflows.
- Integrate multiple and third-party workflows.
- Track and monitor activities.
- Send alert notifications for attempted violations and delays.
- Create alternate routings to cover absences, vacations, and business disruptions.

Electronic Forms and Documents

- Eliminate manual and paper forms and documents.
- Index and classify all electronic documents upon creation.

- Upon creation, create metadata information available for all documents—critical in legal discovery and audits.
- Make it painless to search and recover all related documents—federated document management.
- Enable legally and contractually compliant electronic signatures.
- Ensure a complete information life cycle, including enforced destruction when retention dates are reached.
- Upload electronic forms created offline.
- Enforce version and access controls.

User-Friendly Interfaces

- Create easily understood process flows.
- Show simple graphical representations of workflows.
- Seamlessly integrate to multiple applications.
- Provide online help.
- Access online electronic forms.
- Sign on once to access all accounts and applications.

SURVEY OF CROSS INDUSTRY DEPLOYMENTS OF BPM SOLUTIONS

In October 2006, BPMInsititue.org conducted a web-based survey of "1000s of enterprises across representative samples of selected vertical industries." It included large and small to mid-sized enterprises. Over 75 percent of correspondents had either deployed or were planning on deploying a BPM solution in the next 12 months.[2]

The BPM projects can be characterized as follows:

- About one third of projects were in financial services.
- About one third were planning on spending over $1 million in the current and subsequent years toward a BPM solution.
- At least 70 percent involve human processes.
- About half involve customer and partners outside the firewall.
- About half involve uses not always online.
- About 40 percent involve company confidential information.

Respondents provided the following rationales for their BPM investments:

- Streamline business processes and operating efficiency—about 80 percent.
- Improve visibility and control of processes—about 80 percent.
- Automate manual processes—about 80 percent.
- Improve quality of processes—over 70 percent.
- Improve understanding of current processes—over 70 percent.
- Improve resource allocation and management—over 60 percent.
- Reduce development and maintenance costs—over 60 percent.
- Rapidly deploy new applications—over 60 percent.
- Address compliance requirements—37 percent.
- Provide new revenue opportunities—32 percent.

BENEFITS OF BPM OVER TRADITIONAL PROCESS DEVELOPMENT

Traditional process development is flawed in that they assume business owners always know what they want. They obviously know they have a problem and that they want it fixed, but asking them what an optimized business process should look like is unrealistic for the following reasons:

- Business owners know they have a problem but often are not qualified to define an optimum solution.
- Implementing a solution based on a requirements gathering exercise will rarely provide a best practice and optimized process improvement.
- User acceptance testing is typically the first opportunity for business users to realize the solution they approved is less than optimum.
- Requirements will continue to change as business users begin to use new processes.
- Traditional process improvement projects relied on a waterfall process where all tasks come together after a long development process with little opportunity to modify designs.[3]

The following case studies in financial services provide good examples of the major benefits that can be quickly realized with BPM over traditional approaches to process improvement.

PULTE MORTGAGE CASE STUDY

Derke Miers in *BP Trends* provides a case study of Pulte Mortgage, which sought to improve its customer service through the more timely completion of customer-facing tasks. Without automated workflows, it was difficult to measure processes or spot areas for improvement. By deploying an automated case tracking system, it was much easier to spot areas for process improvements. For instance, it was now possible to track the time to process a mortgage from creation to the point of offer.

By adjusting the credit scoring threshold, managers could shorten or lengthen the process while adjusting their level of risk. This balancing of processing times and risks allowed Pulte to create more dynamic business rules. The rules even included alerts as to when to adjust the automatic credit scoring.[4]

AMERIPRISE FINANCIAL CASE STUDY

Nicole Kealey, group product marketing manager for financial services at Adobe Systems, presented a case study during a November 2006 BPMInsitute.org web-cast to demonstrate the benefits in a BPM solution for a large financial services organization.[5] The objective was to standardize and automate the account-opening

process. The existing process was labor intensive and paper based, with poor version control of documents, and the typically higher risks and costs that come with any manual process. The BPM solution included:

- Electronic forms that enforce policies.
- Ability to capture data both on- and offline, and within and outside the company's firewall.
- Electronic workflow rules, which include a review approval process.
- Document life-cycle management from creation, distribution, and archiving.
- Ability to track and monitor the end-to-end process.
- Provide a transparent audit trail to comply with all applicable regulations.

The benefits included:

- Over a 50 percent reduction in cycle times.
- $5 million in document handling costs.
- Major reductions in the risks inherent in a manual process.
- Major reductions in processing errors and delays.

BPM is typically a successful methodology because it assumes business owners have only a general idea of what they need and these requirements will continue to evolve as process and technologies are deployed and utilized. When replacing inefficient manual processes for risk management with automated workflows and controls, business users will typically realize additional opportunities to streamline processes. Therefore, a BPM approach would seek to apply a Pareto approach: provide the 20 percent of functionality that offers 80 percent of the benefits. As business users use the functionality, they will become much more expert in proposing additional and changed functionality. So, unlike the rigidity of the waterfall approach, the BPM approach encourages feedback and changes in the original designs of a process. This is especially critical with the complex and often convoluted nature of financial risk management.

Exhibit 12.1 is a graphical representation of the traditional straight-line approach to process improvements against a curved line, of BPM and continuous improvement that emphasizes an ongoing monitoring, review, and redesign.

LEAN SIX SIGMA'S SIPOC APPROACH TO BPM

SIPOC is an acronym that stands for suppliers, inputs, processes, outputs, and customers. It is a Six Sigma diagram tool and methodology applied to process improvement projects before the actual work begins. It seeks to identify all applicable events and is usually employed during the measure of a Six Sigma project. (Six Sigma typically uses the DMAIC process improvement approach, which goes through fives phases: define, measure, analyze, improve, and control.) In Six Sigma, suppliers and customers are both internal and external. A customer could be another department, a regulatory agency, a paying customer, and so on.

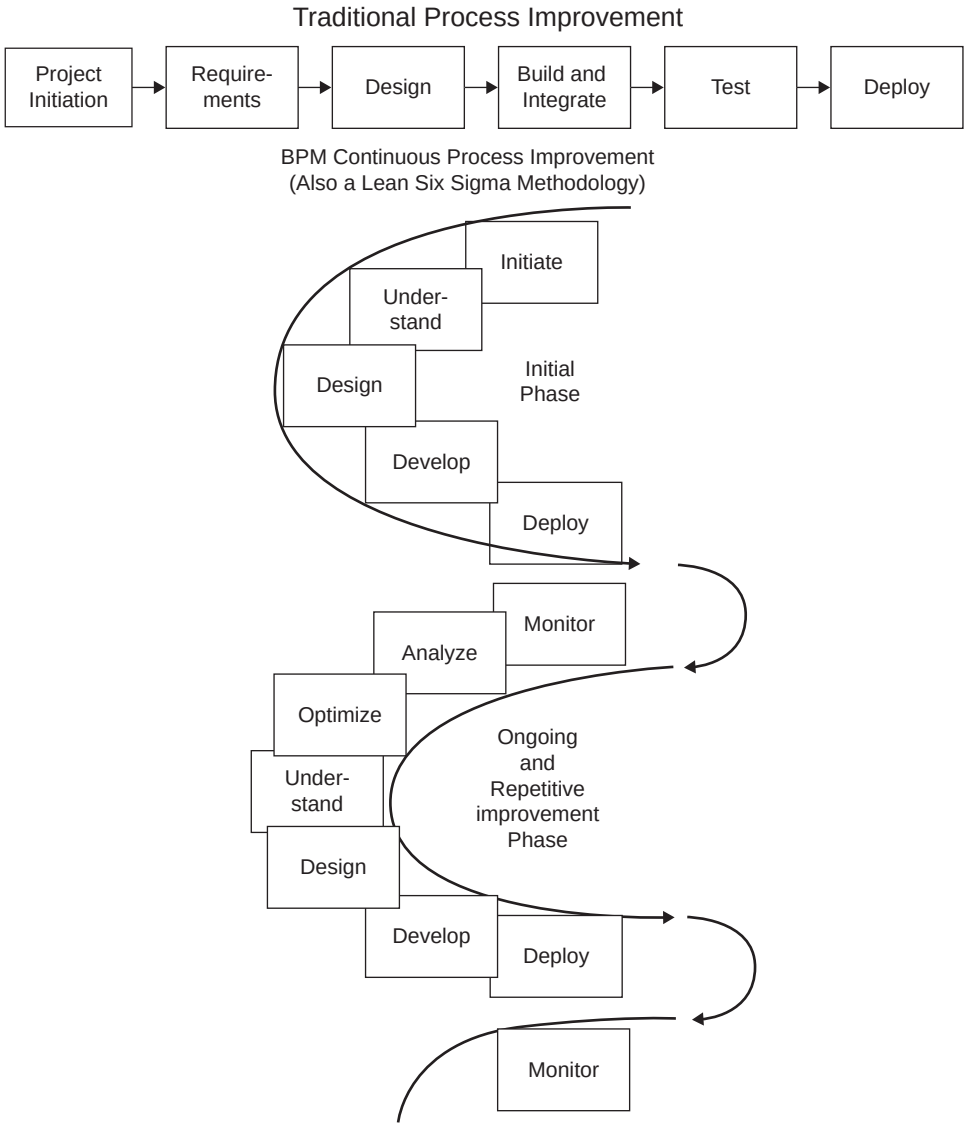


EXHIBIT 12.1 Traditional Process Improvement Projects vs. BPM and Continuous Process Improvement Projects

According to Kerri Simon, writing for the SixSigma.com web site, a SIPOC approach can be especially helpful when the following is not clear in a process:

- Who are the suppliers of the inputs to a process?
- What are the specifications being placed on inputs?
- Who are the actual customers of a process?
- What are all the requirements of these customers?[6]

This can be very useful in financial service risk management due to its complexity and rapid changes. In a simple world, there would be a single supplier input and single customer output for any given process. In controlling risks within the financial services industry, nothing is simple. There are typically multiple regulatory and stakeholder suppliers and customers to any given process. In some cases, the suppliers and customers are not well understood or even known. To add more confusion for a given process, the same entity can be both the supplier and the customer in a process. For example, in creating a purchase order, an external supplier provides the quote, which is one of the inputs and hence is one of the suppliers to the process. The output is a purchase order that the customer sends to the supplier, and hence the supplier is also a customer of the process.

Exhibit 12.2 provides a template for a SIPOC diagram. They are fairly easy to create, and Simon provides the following steps to guide its users to success:

- Identify the outputs of the process.
- Identify the customers who receive the outputs of the process.
- Identify the inputs that the process requires in order to function properly.
- Identify the suppliers of the inputs that are required in the process.
- Optionally, identify the preliminary requirements of the customers, which will be verified during a later step of the DMAIC measurement process.
- Obtain validation and verification from key stakeholders in the project.[7]

CONCLUSION

Exhibit 12.3 is inspired by the work of Forrest W. Breyfogle III. We have extended it to include Lean and show the benefits of the combining the methodologies.[8]

As the above matrix demonstrates, combining Lean Six Sigma with BPM offers many advantages to any organization and can be especially valuable in improving financial risk management:

- *Lean* brings the never ending desire to attack waste of any kind and the realization that is a continuous process, not a project. Lean has proven itself over four decades. It is a philosophy as much as a methodology and is not a slave to a technology solution.
- *Six Sigma* brings proven statistical analysis and problem-solving tools and techniques to bear in the hands of trained, certified, and highly focused professionals. Few business managers, even from the more prestigious business schools, are taught project management, let alone these proven techniques. Six Sigma does not require a burdensome bureaucracy and, like Lean, is not a slave of any technology solution. Much of the work is process driven with simple technology tools.
- *BPM* provides the technology to automate and optimize business processes. BPM is an enabler of Lean and Six Sigma, and in the hands of Lean Six Sigma black belts facilitates the process over manual or disparate systems. BPM permits them to easily use graphical visualization tools to understand, simulate, design, and produce end-to-end workflows that may involve multiple systems, internal resources, suppliers, and customers. BPM also helps to coordinate

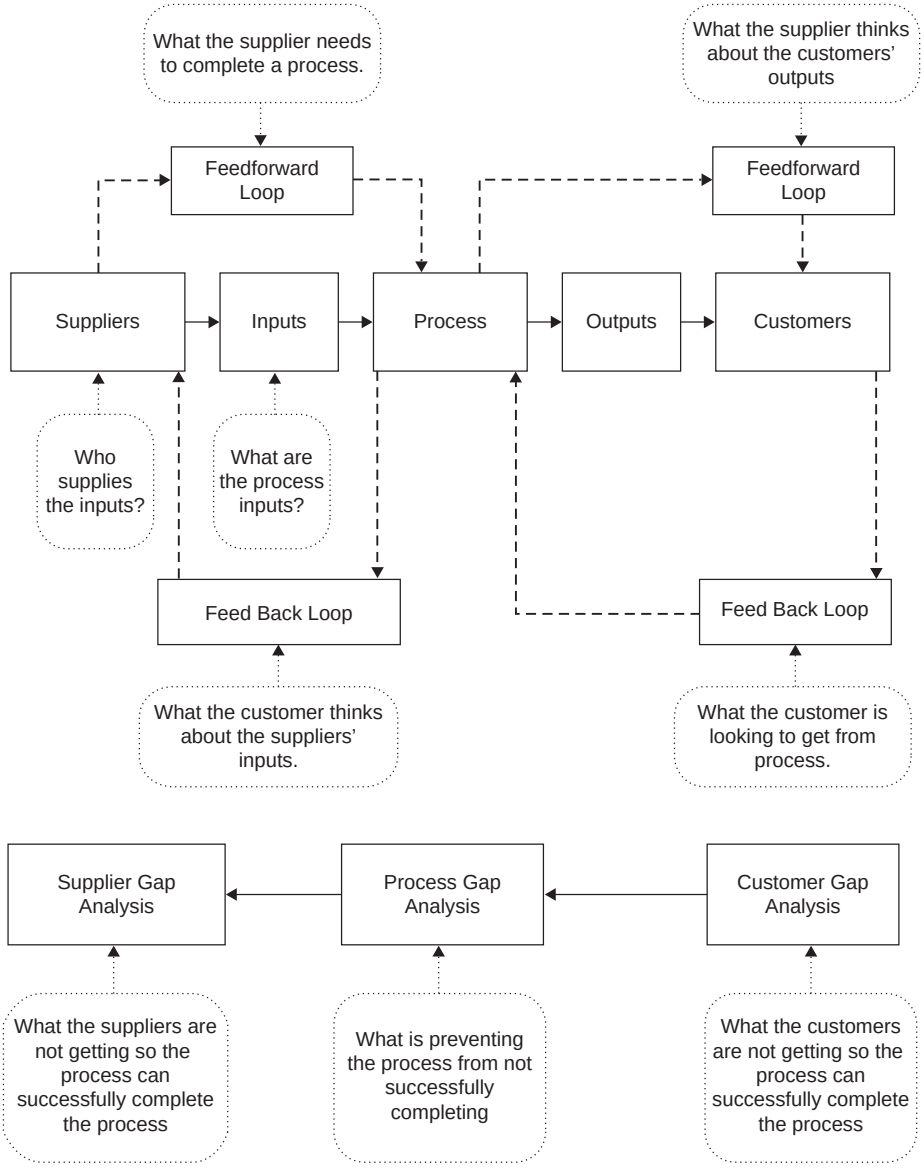


EXHIBIT 12.2 SIPOC Process Map

business processes between people and the data in computer systems by making the data understandable and usable by business users.[9] Once the automated processes are in place, BPM provides the means to monitor, audit, and measure its usage.

Why is this so important in financial risk management? The answer comes from a now famous quote by legendary American bank robber, Willie Horton. When asked why he robbed banks over other businesses, his answer was simple: "That's

	Lean Six Sigma	Business Process Management	Combined Methodology Advantages
Focus	Analysis and Process Improvement	Process Optimization and Automation Technology	Lean Six Sigma's analytical and problem solving processes are enhanced by working with BPM's technology to simulate, design, deploy, monitor, and measure automated processes.
Methodology	Applies a consistent methodology for statistical analysis, process improvement, continuous waste reduction, and problem solving to lower costs and increase service levels to customers.	Typically applies technology tools to optimize and automate processes to assure consistency, shorter cycle times, and lower costs.	Combined they leverage the process expertise of black belts with BPM technology to increase financial returns. BPM provides the graphical dashboards to monitor and measure key performance indicators.
Data	Use statistical tools and analysis of key metrics to identify improvement opportunities.	Access data from enterprise systems to enable statistical analysis.	BPM work flows create a graphically visible and simple to understand source of data to feed Six Sigma's statistical processes.
Design of Process Improvements	Applies root cause analysis to solve problems and improve processes—both manual and automated.	Offers visual graphical design to define process flow improvements—processes which are both man and machine based.	Root cause analysis is facilitated with the visual graphical tools that BPM provides.
Execution of Process Improvements	Creates a plan for process improvements and means to measure their success.	Improved process automated and integration with existing IT investments.	Appling the process improvements to an automated and optimized electronic workflow assures repeatability and reliability in execution. Dashboards provide the needed alerts when out of tolerance conditions are met.
Measurement of Performance	Uses control charts to measure key metrics, provide feedback to sustain the process, and recommend further improvements.	Dashboards measure key process, risk, and performance indicators providing alters when out of tolerance conditions occur.	Migrating from manual and/or Excel-based control charts to fully integrated and hierarchical dashboards enhance measurement, monitoring, and corrective actions.

EXHIBIT 12.3 The Elements of Lean, Six Sigma, and BPM and the Advantages of Combining Them

where the money is." Combining Lean, Six Sigma, and BPM is a great way to balance opportunities and risks. The very nature of financial services provides extraordinary opportunities, but, as the global financial crisis demonstrates, also presents risks that can jeopardize the very existence of firms of all sizes and levels of sophistication.

Combining Lean Six Sigma and BPM should also have the collateral benefits of improving processes and their corresponding information flows in terms of

cycle times, repeatability, costs, transparency, and auditability. An optimized and automated process flow combined with electronically controlled forms provides a transparent audit trail and more timely access to accurate information—all critical in financial services.

Under the new U.S. Audit Standard Number 5 (AS5), which went into effect in 2008, U.S. firms can use BPM technology to benchmark their automated controls and reduce audit costs by up to half. AS5 also permits organizations to use prior year audit results and more heavily rely on internal resources—internal auditors, business owners, and IT professionals. AS5 is bound to have an influence on the emerging International Standards of Audit (ISA), and as such will reward those who combine Lean Six Sigma and BPM. Besides the lower external audit fees, their internal costs of compliance and operations should be reduced as well.

Lean Six Sigma combined with BPM is no panacea or cure-all, but combines three very well vetted methodologies into a commonsense approach. Each has value independently, and when combined can provide significantly greater value. It requires little administrative overhead, can begin with some low-cost pilot projects, and can start showing significant results in three to six months.

NOTES

1. Andrew Spanyi, "BPM Governance," BPMInstiitute.org web site, June 6, 2008, www.bpminstitute.org/articles/article/article/bpm-governance.html.
2. 2006 BPMInstitute.org Member Survey: State of BPM, November 2006, www.bpminstitute.org/roundtables/past-round-table/article/bpm-and-financial-services.html.
3. See Derke Miers, "Getting Past the First BPM Project: Developing a Repeatable BPM Delivery Capability," *BP Trends,* March 2006, www.bptrends.com/publicationfiles/03%2D06WP%2DBPMDeliveryCapability%2DMiers%2Epdf.
4. Ibid.
5. Nicole Kealey, "The State of the Business Process Management Market," BPMInstitute.org webcast, November 26, 2006, www.bpminstitute.org/roundtables/past-round-table/article/bpm-and-financial-services.html.
6. Kerri Simon, "SIPOC Diagram," SixSigma.com web site, www.isixsigma.com/library/content/c010429a.asp.
7. Ibid.
8. Forrest W. Breyfogle III, "Leveraging BPM and Six Sigma," *BP Trends,* October 2004, www.bptrends.com/search.cfm?keyword=Breyfogle&gogo=1&go.x=98&go.y=4.
9. Ibid.

CHAPTER 13

Bayesian Networks for Root Cause Analysis

Deborah Cernauskas, Ph.D.

INTRODUCTION: RISK QUANTIFICATION IN FINANCE

The business world's view of data has changed significantly over the past 10 years. Companies now view data as an asset that, if tapped, can provide enormous value. Accordingly, companies have exponentially increased their demand for data analytics of all kinds. This has subsequently put pressure on the quality demands of the company in all areas of their business, including risk management.

A best-practice approach to operational risk management includes a quality program focused on process control, which is inherently dependent on good data. Although the principles of process control are not industry dependent, the methodology is more prevalent in some industries than others. Process control is widely employed within manufacturing environments and is only recently starting to make inroads into financial firms and financial departments of major corporations. The impetus for process control adoption in financial firms centers on product quality improvement, cost reduction, and customer satisfaction. Whereas quality programs have associated implementation expenses, there are offsetting expense savings from a reduction in rework and an increase in revenue from delighted customers. Quality adds value.

Quality programs have traditionally been the mainstay of manufacturing firms where the application is fairly straightforward. Manufacturing processes are generally well understood. The outputs are discrete units subject to production specifications, which facilitates the identification of quality failures. The application of process control quality programs in finance gives rise to problems not found in manufacturing environments. The processes of financial firms, including banks, institutional investors, and hedge funds, are highly sophisticated and form a complex network for which a measurement system that easily quantifies quality does not exist. Outputs of one process are routinely used as inputs to another. Many of these processes involve large quantities of data from disparate systems, multiple geographic locations, external vendors, and counterparties. Financial processes are difficult to control for many reasons, including a high volume of data processed, the interconnectivity and interdependencies of the processes, and the interweaving of automation with planned human intervention. These obstacles do not suggest that process control quality programs are not viable in financial firms. Quite the contrary.

Financial firms desperately need quality programs because of the complexity of their processing systems coupled with the quick pace of new product development, the high volume of transaction processing, and fierce competition.

Quality problems occur when process variability goes out of control. Process variability results from either random or assignable causes. A process operating with only random causes of variation is considered to be in control and does not require intervention. Assignable causes are due to external elements such as bad data or a software programming error. The presence of assignable causes results in the process's going out of control and requires a root cause analysis to eliminate. Quality-savvy companies approach risk proactively through identification and management.

A critical step in a quality program is root cause analysis to identify process risks and failure points. Common tools used to detect the root cause of a problem are fault trees, fishbone diagrams, and interrelationship diagrams. These graphs are developed using expert knowledge, but are lacking in their ability to define causal links. A new approach to root cause analysis is the application of Bayesian network analysis. Although Bayesian networks have been used for many years to represent complex networks with multiple interacting variables in medical diagnosis and criminal forensics, they are only starting to be applied in process modeling for risk assessment in finance.

This chapter focuses on how to develop a Bayesian network to identify process failure points and uncover causal relationships.

CAUSAL KNOWLEDGE DISCOVERY

There are many viable techniques used for causal knowledge discovery. They include root cause analysis, fishbone diagrams, and fault tree diagrams.

Determining Causal Links

Root cause analysis is a methodology for identifying factors underlying process variation and failures. Direct causes lead to process variation or failure without any other intervening event and will be in close proximity to a failure point. Direct causes are easier to identify in a manufacturing environment than in a financial processing environment, which is comprised of a highly interdependent network of computer programs and databases.

Root cause analysis should probe beyond direct causes. Many problems in a business processing environment occur because a computer application fails to perform properly or is entirely inaccessible. For example, a business cannot issue customer bills when the billing application is down. The direct cause of the failure is a billing application failure, but this information is not sufficient to prevent the failure from reoccurring. It's important to dig deeper and find out the true reason for the application failure, such as a network failure or an application upgrade or patch installation. The true value of a root cause analysis is in identifying the underlying cause of the failure so the business can take steps to remediate.

Common visual methods of root cause analysis include cause-and-effect diagrams or fishbone diagrams and fault trees. These methods help reduce the complexity of interconnect networks through visualization.

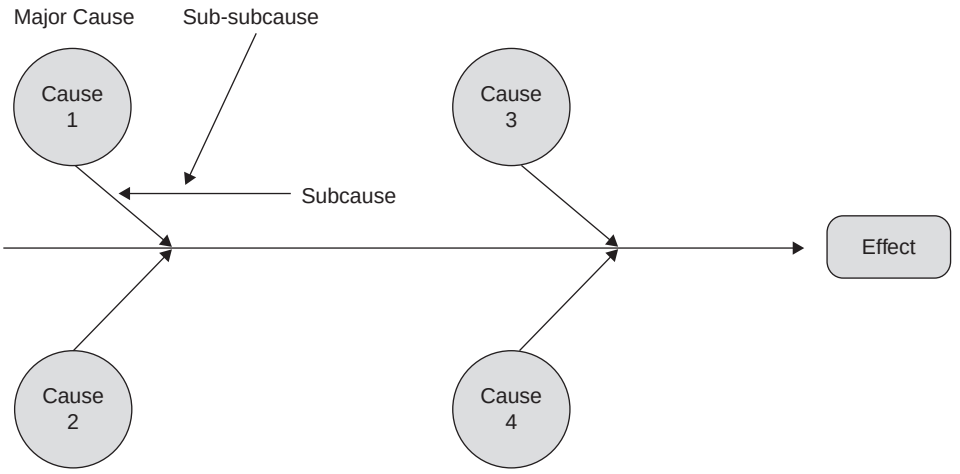


EXHIBIT 13.1 Generic Fishbone Diagram

Fishbone Diagrams

The fishbone diagram has been commonly used to show relationships between problems and potential root causes. Exhibit 13.1 illustrates a generic fishbone diagram. At the far right of the graph is the effect or problem under analysis. The branches and twigs to the left are the possible causes and subcauses of the effect or problem.

Fishbone diagrams aid in the analysis and discovery of potential causes for process failures. A good understanding of the process and problem results in many diagram twigs. A major limitation of the fishbone diagram is that it does not give any indication of the most likely cause of the failure given the available evidence.

Exhibit 13.2 shows a typical vendor payment process for XYZ Corporation. Invoices are submitted to the accounting department from all other departments in the company. An increase in invoice volume due to rapid company growth can easily overload the process, leading to a delay in vendor payments and a loss of prompt payment discounts. The company expects to take advantage of the prompt payment discounts 90 percent of the time. On a monthly basis, XYZ is achieving only a 70 percent rate on available discounts and would like to discover the underlying reason.

Exhibit 13.3 illustrates the vendor payment process as a fishbone diagram. The problem or failure appears at the right side of the diagram. The company wants to take advantage of the vendor prompt payment discount at least 90 percent of the time. The current rate is at 70 percent, which means the company is giving up an easy opportunity to reduce their expenses. The fishbone diagram in Exhibit 13.3 identifies four major causes:

1. An influx of many new vendors.
2. Accounting staffing shortage.
3. A high level of vendor database downtime.
4. Governance process delays.

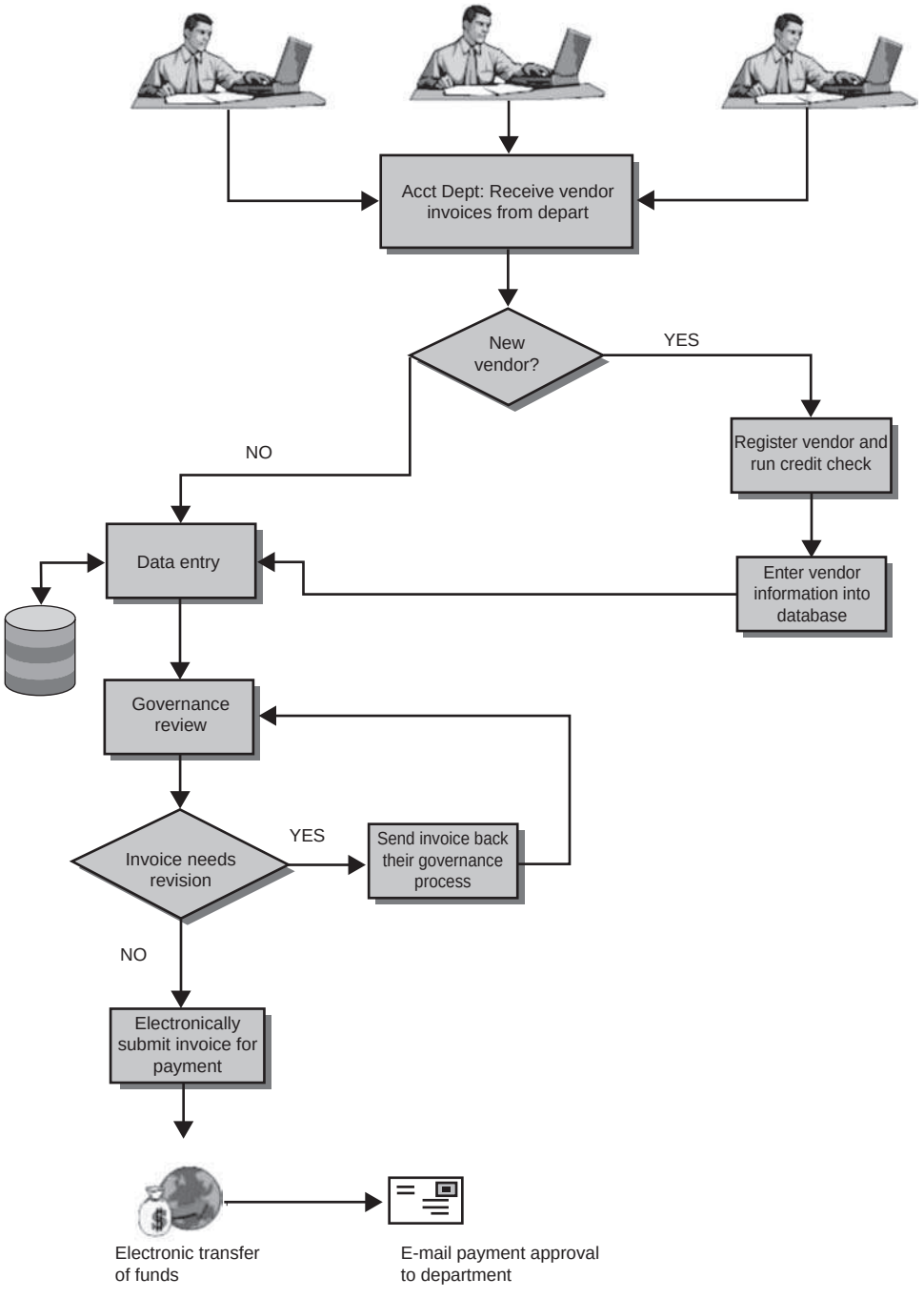


EXHIBIT 13.2 Vendor Prompt Payment Flowchart

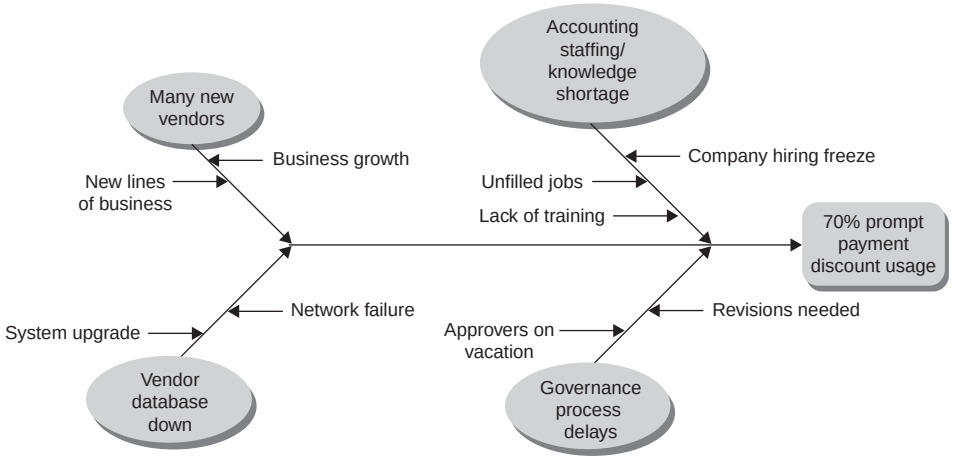


EXHIBIT 13.3 Vendor Prompt Payment Fishbone Diagram

Knowledge of the major cause of the failure is insufficient to take meaningful steps to remediate the process. For example, if it's known that a shortage of knowledgeable accounting staff is delaying the processing of the invoices, it's imperative to know the underlying reason (lack of training, lack of sufficient staff, etc.) in order to take the appropriate corrective action.

Fishbone diagrams are instrumental in thinking through the factors affecting the business process that is not producing adequate output. This diagram will not provide a most probable explanation for the process failure or inadequate process output.

Fault Tree Diagrams

An alternative cause-and-effect analysis is a fault tree analysis (FTA), which employs a top-down approach to determine the possible causes for a failure or bad process outcome. An FTA starts with a failure and works backward through the process to determine the possible root causes. This is in contrast to the bottom-up approach used in failure modes-and-effects analysis, which is a very popular Six Sigma methodology for root cause analysis, as well as in risk planning and management. FTA uses a vertical tree structure to visually show what can go wrong within a process. Fault trees suffer the same limitation as fishbone diagrams—no ability to quantify the most likely root cause (see Exhibit 13.4).

BAYESIAN NETWORKS

A Bayesian network (BN) is a graphical model that integrates the probabilistic relationships between the variables of interest and can be viewed as a probabilistic expert system. BNs are a powerful method for modeling cause-and-effect relationships in complex inference networks and can incorporate quantitative and expert opinion.

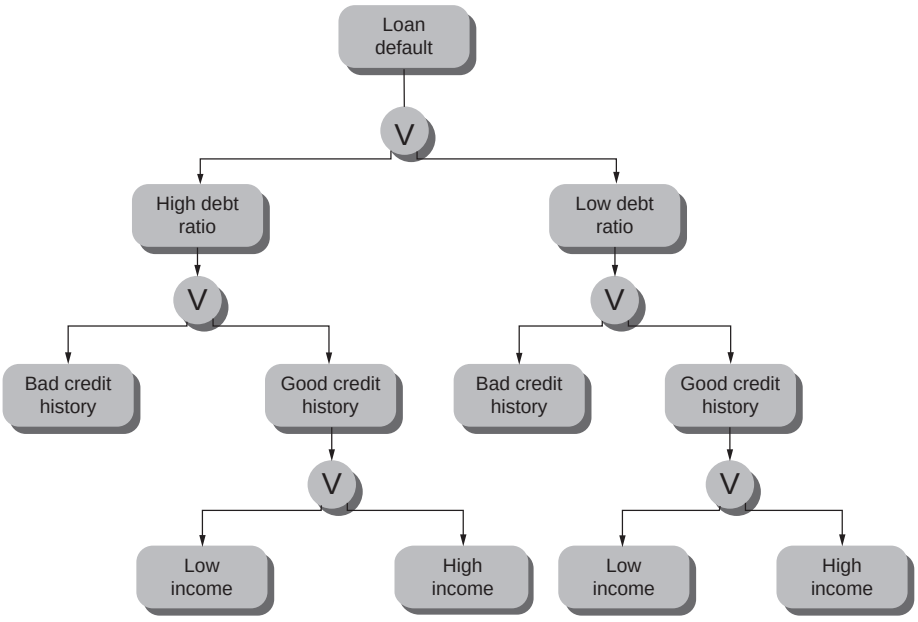


EXHIBIT 13.4 Sample Fault Tree Analysis for Loan Defaults

Introduction to Bayesian Networks

Generally, BNs are represented as graphical networks that capture the probabilistic relationship between variables. BNs are increasingly being used to model systems for which there is incomplete or uncertain information. For example, operational risk quantification can be successfully addressed through BNs.

BN inference is based on the basic law of probability known as Bayes's rule. Given two events, A and B, Bayes's rule can be stated as:

$$P(B_j \mid A) = \frac{P(A \mid B_j) \times P(B_j)}{P(A)} = \frac{P(A \mid B_j) \times P(B_j)}{\sum_{i=1}^{k} P(A \mid B_i) P(B_i)} \tag{13.1}$$

The sample space, S, which is a list of all possible outcomes, can be viewed as the occurrence of event A with all possible values for event B.

To illustrate this concept, suppose you live in Chicago and over time you have determined that during summer it rains 30 percent of the time and that it is cloudy 75 percent of the time (it can be cloudy without rain). The probability of the skies being cloudy given that it's raining is 100 percent, but what is the probability that it rains given that the skies are cloudy? Bayes's rule allows us to compute this probability.

$$P(\text{Rain}|\text{Cloudy}) = \frac{P(\text{Rain})P(\text{Cloudy}|\text{Rain})}{P(\text{Cloudy})} = \frac{0.3 \times 1.0}{0.75} = 0.40 \tag{13.2}$$

A network diagram view of the situation is given in Exhibit 13.5.

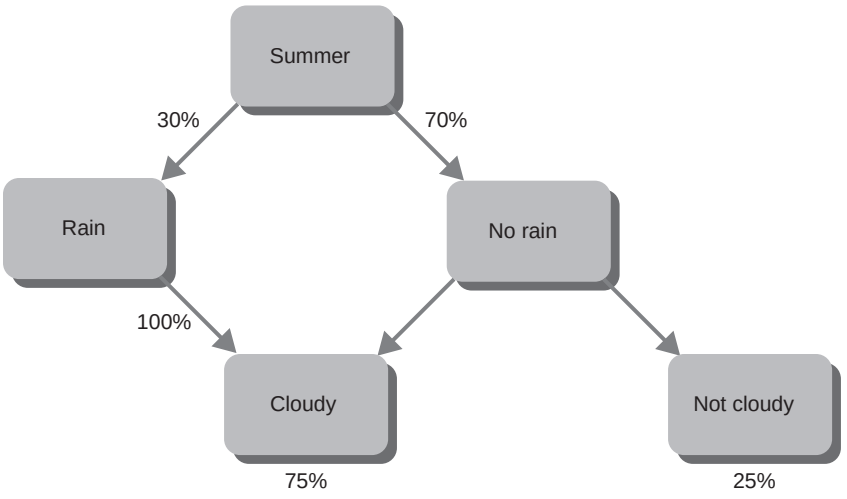


EXHIBIT 13.5 Network Diagram for Rain Example

BN Example

BNs are different from other decision support methodologies in several ways. First, BN models can incorporate expert opinion as input and the models are generally robust to a missing data. Second, a distinct advantage of BNs is the ability to simultaneously model the presence of more than one root cause. Third, BNs allow top-down or bottom-up reasoning. Root cause analysis is performed through bottom-up reasoning, as the network will provide the most likely causes for every effect. Finally, BNs generally relies on a graphical framework to map and quantify the cause and effect relationships between variables. Exhibit 13.6 illustrates a simple Bayesian

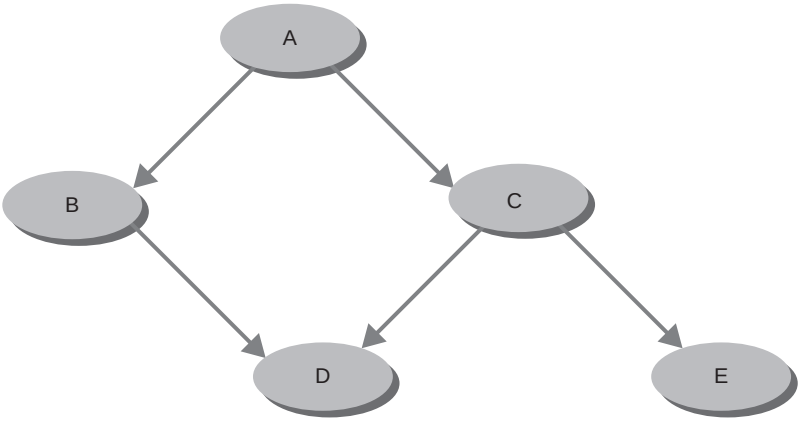


EXHIBIT 13.6 Bayesian Network Example
Source: Richard E. Neapolitan, *Learning Bayesian Networks*, Upper Saddle River, NJ: Prentice Hall, 2004.

network. The nodes represent random variables and are classified as either parent or child. In this simple example, the random variables are assumed to be discrete, taking one of two possible values: "true" or "false."

The network structure is composed of parent and child nodes. Node A is a parent to nodes B and C, and node B is a parent to node D. The arcs in a BN represent causal relationships and the direction of the arrow indicates the direction of influence. Attached to each node are possible states and the probability of being in each state (local probability distribution).

The prior or unconditional probability for node A, P(A), is 0.20. Hence, the probability of not A,P($\bar{A}$), is 0.80. Exhibits 13.7 through 13.10 contain the conditional probabilities for the remaining nodes.

The most probable explanation (MPE) can be found by choosing the set of variables that maximizes P(d|m) where D is the set of possible explanatory variables and M is the manifestation set or evidence set. Suppose in the BN illustrated in Exhibit 13.6 we have m = $\{a_1, e_1\}$ and D = {B,C}. The most probable explanation for m maximizes:

$$P(b_i, c_i|a_1, e_1) \tag{13.3}$$

EXHIBIT 13.7 Conditional Probabilities for Event B Given Event A

	B = 1		B = 0	
Events	**A = 1**	**A = 0**	**A = 1**	**A = 0**
P(B|A)	0.25	0.05	0.75	0.95

EXHIBIT 13.8 Conditional Probabilities for Event C Given Event A

	C = 1		C = 0	
Events	**A = 1**	**A = 0**	**A = 1**	**A = 0**
P(C|A)	0.10	0.3	0.9	0.70

EXHIBIT 13.9 Conditional Probabilities for Event D Given Events B and C

	D = 1				D = 0			
	C = 1		C = 0		C = 1		C = 0	
Events	**B = 1**	**B = 0**	**B = 1**	**B = 0**	**B = 1**	**B = 0**	**B = 1**	**B = 0**
P(D|B, C)	0.8333	0.50	0.333	0.1972	0.16.67	0.50	0.6667	0.8028

EXHIBIT 13.10 Conditional Probabilities for Event E Given Event C

	E = 1		E = 0	
Events	**C = 1**	**C = 0**	**C = 1**	**C = 0**
P(E|C)	0.3077	0.8243	0.6923	0.1757

To find the MPE we need to compute the following four conditional probabilities:

$$\begin{aligned} P(b_1, c_1|a_1, e_1) &= P(b_1|c_1, a_1, e_1)P(c_1|(a_1, e_1) \\ P(b_1, c_0|a_1, e_1) &= P(b_1|c_0, a_1, e_1)P(c_0|a_1, e_1) \\ P(b_0, c_1|a_1, e_1) &= P(b_0|c_1, a_1, e_1)P(c_1|a_1, e_1) \\ P(b_0, c_0|a_1, e_1) &= P(b_0|c_0, a_1, e_1)P(c_0|a_1, e_1) \end{aligned} \tag{13.4}$$

For the simple BN the above methodology works adequately. As the size of the BN grows either through the number of variables or the number of possible outcomes for each variable, the number of conditional probabilities that need to be evaluated grows exponentially and more efficient methodologies will be required.

CONCLUSION

Bayesian network analysis provides a new approach to root cause analysis for financial business process failure determination. BN analysis is a robust methodology that allows the blending of expert opinion with empirical data. This is particularly important when modeling business processes for which empirical data may be sparse or missing.

The benefit of Bayesian network models over the standard Six Sigma methods is the assignment of empirically or expert based probabilities, which allows the analyst to determine the most likely cause of failure. Fishbone and fault tree diagrams are good vehicles to use to think through the drivers of a process, though neither diagram leads the analyst to the most probable explanation for the process failure. The availability of commercially available software will facilitate the use of the methodology. Bayesian network analysis is a necessary addition to the Six Sigma toolbox.

BIBLIOGRAPHY

Neapolitan, Richard E. *Learning Bayesian Networks* Upper Saddle River, NJ: Prentice Hall, 2004.

Neil, Martin, Norman Fenton, and Lars Nielson. "Building Large-Scale Bayesian Networks," *The Knowledge Engineer Review* 15(3) (2000): 257–284.

Pourret, Olivier, Patrick Naim, and Bruce Marcot. *Bayesian Networks: A Practical Guide to Applications*. Hoboken, NJ: John Wiley & Sons, 2008.

Taroni, Franco, Colin Aitkem, Paolo Garbolino, and Alex Biedermann. *Bayesian Networks and Probabilistic Inference in Forensic Science*. Hoboken, NJ: John Wiley & Sons, 2006.

CHAPTER 14

Analytics: Secrets to Deriving Business Value and Insights out of Information

Ying Chen, Ph.D.

ABSTRACT

The volume of information and the speed of its spread create significant financial risks as well as opportunities to businesses. Those who are able to digest such information effectively and timely can derive great insights and gain high business values while reducing their exposure to financial risks. However, if such information is not leveraged, corporations may create significant financial risks for themselves in many different dimensions, such as reputational or operational risks, brand stewardship and image, and competitiveness. Advanced analytics technologies, such as text mining and data mining, show great promise in enabling businesses to effectively utilize vast amount of information for business insights and value. In this chapter, we first provide a brief historical analysis and overview of information-based technology and services trends in the past 40 years. We indicate that our generation is an analytics generation and our future generations will be even more so. We then present a set of key analytics technologies, especially in the text analytics space, that have emerged in the past decade. We show that such technologies can be applied and bundled into specific solutions to tackle specific business problems. In particular, we describe a social media mining solution built on top of such technologies, which mines social media content such as blogs, message boards, news, and Web content for protecting enterprise reputations, and gaining brand, market, and consumer insights. Finally, we highlight several emerging areas in information analytics and how they might impact businesses' risk management in the future, including social media analytics and Web-mining technologies and social network analytics.

This chapter would not have been possible without the work done by researchers and engineers in IBM Research, especially Scott Spangler, Jeffrey Kreulen, Larry Proctor, Bin He, Amit Behal, Ana Lelescu, and many others who helped in the creation of a number of analytics technologies described in this chapter. My sincere gratitude goes to all of you.

INTRODUCTION

Today's businesses increasingly rely on a vast amount of information. Yet effective use of information is becoming more and more difficult. The sheer *volume* and *diversity* of the information represents one of the major roadblocks. Today, corporations not only need to leverage corporate internally generated information in all forms, including structured data such as financial information and unstructured data such as documentation, manuals, e-mails, instant messages, papers and reports, they also need to pay significant attention to external consumer or community-generated media (CGM) information such as blogs, online forums/message boards, news, and Web contents. This is because today's corporations are far more "naked" than before.[1] The ecosystem of a given corporation is complex and dynamic. They often have significant influences over companies' directions. They cannot be ignored.

Corporate internal information provides corporate internal operations insights. Yet corporations are rarely controlled solely by executives or board members in today's well connected world. External communities, governments, nonprofit organizations, consumers, customers, and many other global entities represent a stakeholder web around the company.[2] They have significant influence on corporations' directions and success. CGM information may capture key insights from many of such stakeholders of the company. When effectively utilized, CGM information coupled with corporate internal content can lead to game-changing insights, such as identifying unknown market opportunities well ahead of competition, changing the dynamics of the corporate ecosystem to enhance corporate image and reputation, reaching out to communities that are not aware of corporate product and services, or removing the significant, but latent, threats to the corporation by identifying them early and proactively.

Clearly, dealing with information overload is nontrivial. A spectrum of approaches is practiced widely to date, ranging from *ignoring them by and large*, to *making small effort*, and to *performing deep and thorough analytics*. Davenport et al.,[3] indicated that only those who painstakingly make data available for analytics and proactively engage in analytics in all aspects of business can compete well and lead in the marketplace today.

Although the volume of information and how fast it travels today may suggest that one might need to ignore a significant amount of it to avoid getting drowned. Hence, the "ignore" approach has some wisdom to it. Yet, this is true only if the corporation is very wise and fortunate in selecting just the right information at the right time to make decisions. Obviously, this approach is not only extremely hard to work in the first place, but also cannot last long even if it seems to work for a certain period of time. The second approach, where corporations are only willing to make a marginal effort, will only see marginal results. Deep insights and high values require corporations to put serious efforts and persistence into investing in analytics in all its dimensions, including tools, technologies, and methodologies, hiring or training staff to patiently work with data and tools to derive insights.

The mentality of "simply Google-searching things and expecting the right answers to return instantaneously" will not work. This is true no matter how advanced the analytics technologies and tools might become. Ultimately, it is the human beings' interpretation of the insights surfaced by the tools that can make a difference.

Given the complexity of the data and ecosystem, no substantial insights are possible without tenacious efforts from human beings.

In section 2 of this chapter, we provide a historical view on the evolution of information based technologies and services. We indicate that our current generation needs to be an analytical and insight-driven generation. Without analytics, we will lose sight of the right directions and be overwhelmed by the irrelevant. This is even more so for the future generations.

In section 3, we lay out a landscape of the analytics technologies that have become available in the past decade. We especially focus on text mining related technologies as it is recent and has major technical challenges still to be addressed. We then demonstrate that such analytics technologies can be successfully applied in real-world situations by illustrating a domain-specific analytics application that mines the CGM content to derive insights for brand and reputation protection by detecting alarming signals early.

Section 4 examines the emerging areas in analytics technologies, specifically around social media analytics, Web 2.0, and social networks, aside from traditional corporation business intelligence. We argue that the dynamic network of the stakeholders of a corporation mandate acute understanding of its environment and leverage such network for the benefits of their own and the world it lives in.

Finally, we draw conclusions in section 5 and highlight the key challenges in today's analytics technologies and their future directions.

INFORMATION TECHNOLOGY AND SERVICE EVOLUTION

In the 1970s, computer systems were born and started to impact businesses and people's lives in a limited fashion. From information technology perspective, much of the focus has been on the speed and the capacity of the data storage. Megabyte of storage was a luxury at the time. There was a desperate need to scale up the *capacity and speed* of information access for such computing technologies to be widespread.

From the mid-1980s to 1990s, technology advances made speed and capacity much less an issue. Gigabyte storage became commonplace. New technologies such as relational database management systems (RDBMS) and file systems emerged. They started to enable fast transaction processing and business operation automation. In the late 1990s and early 2000s, content management systems arrived. They were often integrated with different business processes to automate all aspects of business operations that involve document processing. Clearly, that generation was all about *automation and business process integration*.

In the late 1990s and 2000s, the emergence of the Web and further advanced in information technologies changed the world completely. Information storage capacity and speed of access have exponentially advanced. Even terabyte storage is easily affordable by all enterprises and many individuals. Everyone has easy access to all kinds of information at all times through the Internet. The concepts of business intelligence (BI) started to come about. Data mining has evolved from academic exercises to real-world practices by mining information stored in RDBMS (called "structured" information) to find hidden associations (e.g., the relationship between the sales of a given product and buyer's gender, location of purchase, and time of purchase).

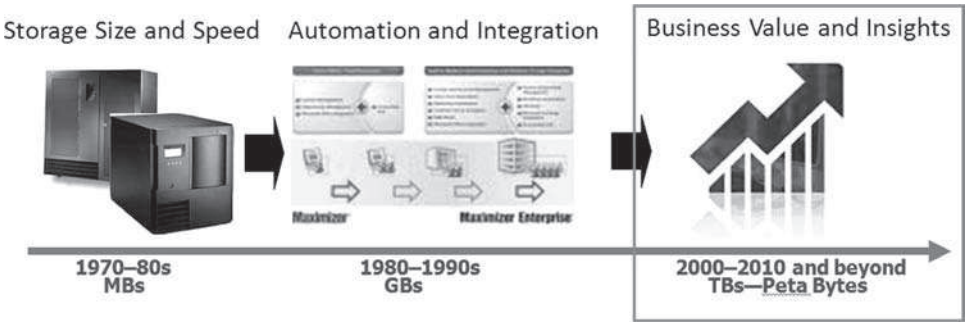


EXHIBIT 14.1 Information Technologies and Services Evolution

Today, online analytical processing (OLAP) and BI solutions are addressing challenges in many aspects of business, ranging from customer relationship management (CRM) to financial performance analytics and optimization.[4,5,6,7,8]

However, structured information accounts for only a small fraction of the total information population. The prominence of unstructured data such as e-mails, instant messages, and various forms of documents (e.g., Word, PowerPoint, PDF, Web) demands even more advanced information analytics approaches. In the past decade or so, a wide variety of text mining techniques for unstructured data have been developed, including smart information retrieval,[9,10] natural language processing (NLP) to extract semantic entities out of text (also called "annotation"),[11,12] clustering, classification, and taxonomy generation[13,14] to analyze large body of related textual information (also called "corpus") through autocategorization. Today, although less prominent than traditional BI and OLAP technologies, text-mining technologies start to provide key insights in many business functions (e.g., CRM for customer satisfaction analysis using customer survey comments, contact center call log analytics to allow efficient problem resolution, intellectual property [IP] for patent portfolio analysis and licensing purposes, and Web and social media analytics for market and consumer insights and brand image and reputation protection).

Clearly, our generation is abundant with widely available information. The key to the successes of today's businesses lies in how one can leverage such information to derive critical *insights* and turn them into *values* to the business while reducing its financial risk exposure. To do so, corporations must embrace analytics all around—leveraging analytics tools and technologies, training and hiring employees to acquire analytical skills, and investing in development of analytics technologies. Exhibit 14.1 shows such an evolution. We believe that the future generation will require even more analytics, given the ever-increasing amount of information in all forms.

INFORMATION ANALYTICS TECHNOLOGY LANDSCAPE

In this section, we lay out a landscape of information analytics technologies existent today. In particular, the information analytics technologies can be roughly summarized into two categories: data mining, which focuses on structured data stored in

RDBMS, and text mining, which mines unstructured text. Other mining techniques such as video or audio mining also exist. But they are far less mature than data- and text-mining technologies. Hence, this chapter primarily focuses on data- and text-mining technologies. Between data-mining and text-mining technologies, text mining is currently drawing significant attention because of the volume of unstructured text and the vast CGM content in social media space such as blogs, message boards, news, and Web.

This section specifically provides an overview of text-mining technologies and its applications in real-world cases. We illustrate such technologies from an architectural viewpoint. We show what key component technologies are required to compose a real-world analytics solution and the methodologies required for operating such components to address real-world problems. We describe a specific application of such analytics technologies for social media mining for the purpose of corporate brand and reputation risk protection. We also describe data-mining technologies and text-mining technologies. We then present real-world use cases of such analytics technologies.

Data Mining

As the relational databases grow in size and table relationships become increasingly complex, data-mining techniques emerged. Data mining aims at finding hidden patterns in data and relationships by mining large relationships in databases. In general, any real-world data mining solution requires three key technology suites:

1. *Extract, transform, and load (ETL) solutions* for data processing, cleansing, and data warehouse building.
2. *Data mining analytics algorithms* that identify hidden patterns and relationships.
3. *Visualization and reporting front-end technologies* that allow end users to quickly review analytics results and compose analytical reports.

ETL solutions became available in late 1990s and received significant attention in the marketplace in the 2000s. Today, many vendors offer ETL solutions[15,16,17] to enable flexible, scalable, and efficient data loads into the appropriate data schemas that supports data mining and BI operations, such as OLAP rollups and slice-and-dice operations. Typical ETL's target data models are *star* and *snowflake* schemas[18] which are designed to enable fast rollup and aggregate operations for large quantity of data in RDBMS.

Besides ETL and data warehousing, data-mining algorithms were a hot topic in academia in the 1990s, and they advanced significantly in the early 2000s. Data-mining techniques often center on machine learning and artificial intelligence, such as neural networks, decision trees, naïve bays, and other predictive modeling techniques.[19,20,21,22] Such techniques often are aimed at finding unknown but significant patterns through computer-aided algorithms. Today, such techniques have been applied in many real-world situations such as understanding consumer buying behavior (e.g., demographics of products or services, forecasting for retail inventory management, financial analysis, and predictive models for market outlook).

Even with mining algorithms and computer help, the end insights can be spotted only by analytical-minded human beings. To facilitate easy interpretation and interaction of analytical results generated by the tools and algorithms, data-mining and BI vendors have also been devising significant visualization and reporting techniques. Often, the underlying slice-and-dice operations may result in different graphical views of analytical results that are designed for different purposes. In summary, although the core technologies of data mining are around analytics algorithms, other techniques such as ETL and visualization and reporting are critical to make mining accessible and practical for real-world usage.

Text Mining

Text-mining technologies are much more challenging than traditional data mining, due to their being *unstructured*. Although the key technology suites around text mining can also be clustered into the three categories listed for data mining (i.e., ETL and data warehousing, text-mining algorithms, and visualization and user interaction) each of these three suites of technologies requires significant innovations when dealing with unstructured data. In the following sections, we will highlight key challenges and technologies devised in this space. We describe them by following through the three layers of the technologies in detail. Some example systems that contain such technologies can be found in the work of Behal et al.[23,24] Exhibit 14.2 shows a sample end-to-end analytics system architecture.

- A generic ETL (GETL) engine continuously processes information in structured and unstructured forms and creates an information warehouse as shown on the middle of Exhibit 14.2.
- An analytics engine applies text mining and data mining techniques to derive insights.
- A visualization and user interaction component that presents the analytical results in an easy-to-understand fashion.

ETL Processing ETL for unstructured data differs significantly from those that are designed for structured data. This is mainly due to two key factors: First, unstructured data is significantly larger than structured data. ETL processing for text must scale in both speed and volume. Here, unstructured data also include semistructured data such as extensible markup language (XML). Second, unstructured data require special extraction and transformation processing. For example, web content often contains significant duplicate information. To cleanse such data, duplicate detection and elimination might have to be plugged in for Web content. Without such processing, analytics may not be effective. Often, the blog and new feeds may come in an XML format. The XML tags may contain structured fields, such as the URL of the blog entries and news and the publish dates. Such information may be valuable for analysis.

To address such issues, a GETL framework is needed. Such a GETL can extract and transform data from semistructured or unstructured data into a standard formats via an extraction and transformation framework. For instance, different date/time formats may be transformed into a standard *mm-dd-yy:hour-minute-second* format.

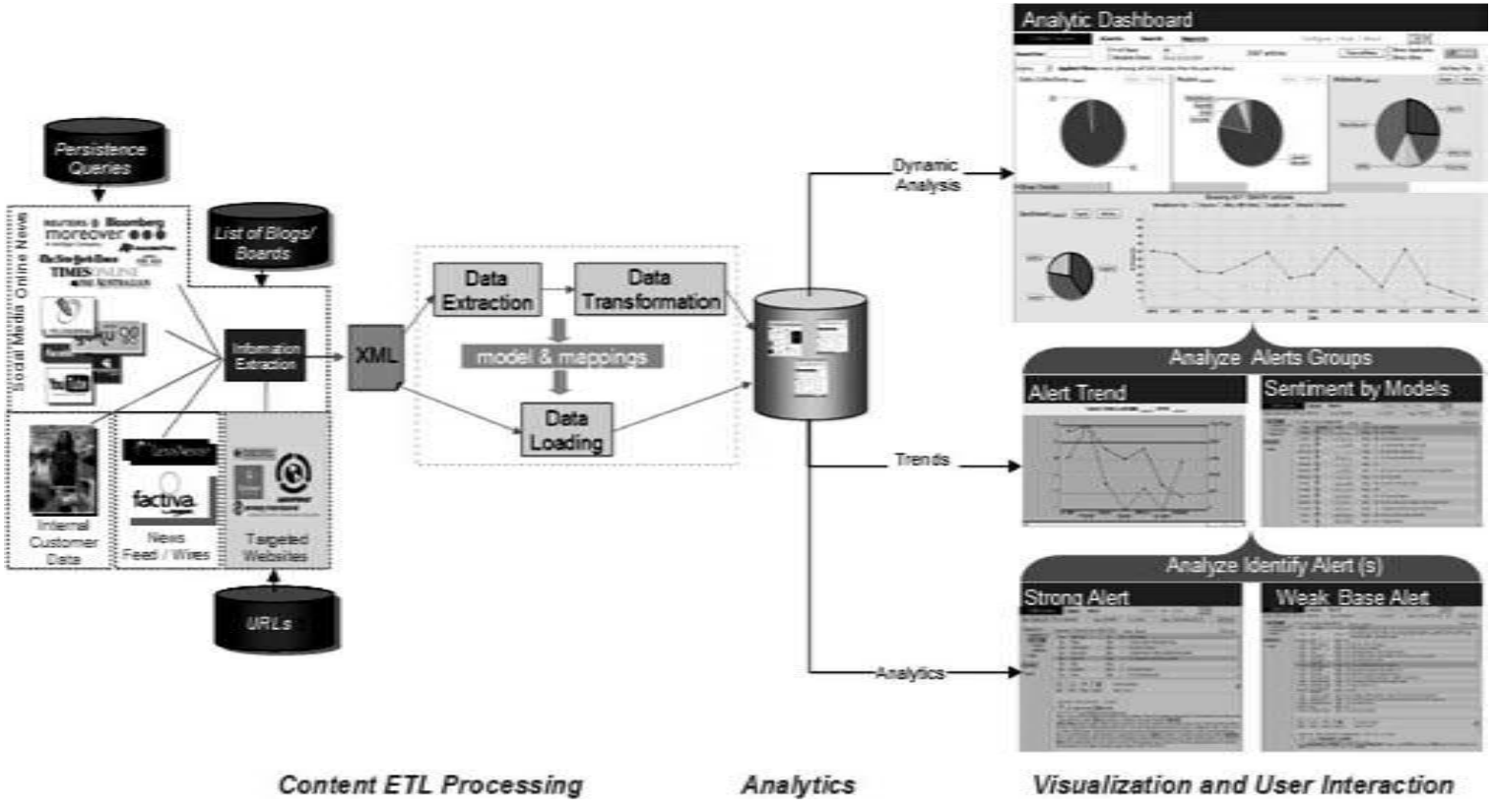


EXHIBIT 14.2 A Sample System Architecture and Components for an End-to-End Analytics Solution

Users can also define their own extraction and transformation functions in GETL, including text annotators that extract semantic entities out of unstructured text[25] or deduplication as described earlier. Once extracted and transformed, GETL loads the data into a target data warehouse. The warehouse can use standard schema such as *star* and *snowflake* schemas or custom schemas to enable traditional OLAP queries and data-mining operations aside from text mining.[26]

Text Analytics Text-mining technologies can be roughly broken down into three categories:

1. *Search-related technologies,* which aim at leveraging appropriate indexing and search algorithms to retrieve relevant information that users are looking for from a corpus of documents.
2. *Smart and semantic information extraction* (also called *annotation*) to extract semantic concepts out of text, such as human names, corporation names, addresses, and phone numbers.
3. *Document collection level analytics* that aims at understanding large corpus of information such as text clustering and classification. This often deals with deriving different taxonomies to allow users to examine data from different angles and correlating them in some meaningful manner. Such taxonomies can be considered as a different "lens" to the data set. Often, deep insights are found at the intersecting relationships of these taxonomies.

To address real-world problems, a combination of these techniques must be used. Many of these techniques also leveraged a combination of machine learning, statistics, artificial intelligence, and NLP techniques.

In general, text analytics solutions have a mission to discover insights that are hidden and nonobvious, under a hypothesis that data could tell the truth or provide some wisdom that one would not be able to gain otherwise. Search solutions such as Google, however, are used to find things that are already known. Search is best when one is to validate certain observations that can be easily validated based on a small set of results returned by search queries. For example, when a Google search returns results, in most of the cases, users either know right away that they have found what they are looking for because the top few pages contain the relevant answers, or they give up because no matter what queries they construct, they may not find the right answers.

Text analytics is best suited when one is willing to "listen to" what the data has to say and willing to follow and interact with what is discovered by the data to reach certain conclusions, which are often unexpected, such as uncovering a market that is unknown before. Because of the discovery nature, users will gain insights only through an iterative analytical process in working with the data and tools, as opposed to instantaneous search results. In short, text analytics is much more data-driven, discovery-centric, and iterative in nature, while search is driven by human knowledge and presumptions and is much more instantaneous. Although both kinds of technologies (i.e., search and analytics) have their values to corporations, we argue that today it is analytics that would give an edge to corporations, because the analytical insights are often game changing and differentiating, while search is commonplace and business as usual.

In general, analytics technologies may be organized into three key aspects:

1. Explore
2. Understand
3. Analyze

Typically, an analysis process often starts with an initial *exploratory phase*, in which a user queries, selects, and extracts information about an area of interest from the information repository. For example, one may want to understand what people are saying about green information technology (IT) from a warehouse that is constructed from a set of technology and IT-centric blog sites. A search query can be used to retrieve all documents that contain the words *green IT* or *green technologies*. These documents are referred to as *seed documents*. They represent a universe of the documents that might be relevant to the subject of analysis. They do not need to be extremely precise to begin with because other analytics tools will facilitate the filtering out of irrelevant information or expansion of additional data if needed. However, the resulting document set cannot be too generic, to avoid diluting the analytics results with irrelevant information. Too specific a query may result in missing relevant information. The main goal of this explore step is to find a sufficient set of relevant data that can be used by the subsequent analysis steps to derive insights. The seed documents can be produced as any combination of structured and unstructured field searches.

Sometimes the overall analytical process is iterative. So it is possible that the initial data set may be too broad or too narrow; the subsequent analysis steps may suggest ways to narrow the scope by providing relevant keywords to search for or to broaden the query in some form, such as looking for documents that are similar to the retrieved data set (called "nearest neighbor search" or "similarity search")[27] to expand the data set.

Annotation techniques can be leveraged as well to allow one to search documents that contain specific semantic entities, such as human names, corporation names, or countries. For example, if a country annotation exists, one can use such an annotator to extract all country names from the text. Such annotation results can be stored back into the data warehouse as part of the structured dimension. During the explore phase, one can extract all documents that mentioned a specific country (e.g., all documents that talk about green IT and the United States).

Annotation is a general technique to understand semantic entities embedded in the unstructured text. An annotation step can be introduced in many places throughout the analytics process. One can apply a human name annotator at the ETL process phase to extract the human names from the documents and populate a human-name dimension as any other structured fields. Or an annotator can be applied after the explore phase to extract semantic entities on the seed document set as described earlier.

Annotators can be built in many ways: a dictionary-based annotator checks the words and phrases in the text against a given dictionary that contains all the terms representing a target semantic entity. For instance, an English human name dictionary contains all possible combinations of human names in English. A word that matches a dictionary term suggests that it is highly likely that it is the target semantic entity. For instance, if "John Smith" is found in the human name dictionary, then it is considered

as a human name. There are also linguistic rule-based annotators which use a set of rules to extract semantic entities; and methods that combine two or more above techniques to build annotators, such as a combination of dictionary and rule-based annotators. Other techniques such as machine learning may also be used to construct annotators based on a set of labeled documents, such as conditional random field and hidden Markov models.[28,29] For the purpose of this chapter, we do not describe the detailed algorithms for developing annotators for annotation techniques.[30] Instead, we show that annotation is yet another text analytics approach that can be used to help understand a given document collection.

The explore phase may produce an unwieldy amount of output for the analyst to sort through and comprehend manually. To aid in comprehension and further refinement of the output, an *understand phase* is carried out. The understand phase focuses on building various taxonomies for users to understand the document corpus from different angles. Taxonomies are critical to human understanding of the large data corpus. Intuitively, the most direct ways for human beings to understand large quantities of information is through categorizations, and maybe categorization by many different ways. In this chapter, we call these categories *taxonomies*. These different taxonomies serve as different lenses for analysts to gain understanding of the document set. Among all the ways of categorizing data, the most important one is "nature classification," which creates categories of documents mainly based on the unstructured text in the documents. If the documents naturally form clusters, then they can be bundled into classes in taxonomy. Such a clustering technique produces a taxonomy that describes what this document set is about without having to read every single document in the set.

Such content-driven taxonomy generation techniques are critical because, even if the documents are already classified by some predefined structured fields, such structured fields may not accurately reflect what content says. This is especially true if the documents discuss an emerging topic that does not fit into any predefined topic category. Machine-aided analytics based on the document content may be much more truthful and accurate in reflecting the real insights. Furthermore, the natural taxonomy can be especially interesting as unexpected categories may surface and bring new insights that cannot be discovered otherwise.

Many methods can be used to generate such taxonomies under different conditions. A popular statistical method is vector space model. Under such a model, each document in the seed document set is represented by a numeric vector that corresponds to its words, phrases, and structured information; then the system can classify the documents into appropriate categories (also called *document clusters* or *classes*) using a clustering technology, where each document cluster represents a set of concepts common to the documents in that cluster. For instance, a green IT taxonomy might contain a cluster on "data center" or "data management" if such topics are prominent and naturally form clusters. Spangler, Kuelen, and Modha describe the details of several such clustering algorithms for taxonomy generation.[31,32]

Other ways of categorizing documents deal with classification based on either a given set of taxonomies or based on the structured fields for the documents. It is not uncommon for corporations to enforce predefined taxonomies on their corporation documents simply because such taxonomies may map to an organizational structure. For instance, the corporate documents may be classified by industries, organizations, or geographical locations. To map documents to such given taxonomies,

classification models can be built to automatically classify documents into such predefined taxonomies. Sebastiani highlights a key set of machine-enabled classification technologies.[33] Many such classification technologies are critical in addressing today's risk and compliance issues faced by corporations. Large corporations are often mandated to manage their corporate documents based on regulatory compliance-driven requirements or taxonomies. In the past, most of the documents were not labeled according to such taxonomies. Classification technologies show great promise for the auto-classification of corporate documents which enable proper document control and management. A real world analytics solution may use a combination of such classification, annotation, and natural classification/taxonomy generation techniques enabling users to create a different lens to look into data and hence gain comprehensive and global understanding about the data set.

Although taxonomies are useful by themselves, cross-taxonomy analysis can gain new perspectives and discover hidden patterns and relationships that lie in the intersections of taxonomies. This is the *analyze phase,* in which co-occurrence analysis methodologies are used to compare any two or more taxonomies. For instance, one can compare the natural taxonomy with a structured dimension such as country to identify relationships in terms of affinity or hidden associations (e.g., which country is highly associated with what kinds of green IT technologies). Furthermore, a trend analysis could be used to understand the changes of such relationships and patterns over time. The co-occurrence is often computed through certain affinity measures such as a statistical Chi-Square test.[34]

In summary, the key unstructured analytics techniques include:

- Search-related technologies.
- Taxonomy generation, including clustering, classification, and annotation techniques.
- Relationship analysis such as affinity analysis and network analysis.

Underneath such a broad categorization, one may use specific techniques to tackle specific problem areas, such as statistical based analysis of words and phrases in the documents, or machine learning methods, or NLP methods.

Visualization and User Interaction Although the analytics techniques and workflow described earlier have logical flow in them, each of these steps may require significant user interaction to be effective in a process that is often iterative. This is because the machine-generated results may lack human domain knowledge and inputs. For example, certain domain-specific words might need to be included into the dictionary manually for the machine to take them into account for special consideration. Otherwise, they may not receive the deserved attention. Interactive taxonomy editing is often the most critical step in generating a meaningful taxonomy that can accurately represent the content as well as human knowledge. Taxonomy editing allows the analyst to refine machine generated clusters for the interactive taxonomy generation techniques). Such user interaction is rarely required for structured data mining.[35,36] Yet they are unique and critical for analyzing unstructured text.

In addition to editing, an analytics solution should also allow the user to save intermediate analytical results for later refinement or reuse. For instance, one can

select a category for further analysis. Additional taxonomies and relationship analysis can be generated on a specific class alone for refinement. This process can be repeated as many times as the user may wish. The overall analysis approach allows the user to zoom in and out between the global perspective created by the high-level taxonomies and co-occurrence analysis as well as detailed perspectives created by subsequent iterations on individual categories of documents. In summary, such iterative analytical approach helps the analyst to understand various aspects of an informational set. The overall process is by no means instantaneous, but it can be extremely enlightening when certain insights are reached.

To ensure flexible user interactions and high-quality insights, different visualization techniques and user interaction techniques are employed. For example, under each type of analysis, different visualization representations are provided wherever appropriate to help the user easily identify and understand the insights. The proper visualization allows users to pinpoint the appropriate insights easily. For instance, multiple visualization representations of a taxonomy view are presented: a list view of classes in the taxonomy is a straightforward way to summarize what is in taxonomy. A table view is more appropriate if additional information such as how many documents there are in each class and various statistics associated with the class are presented.[37]

In summary, a real-world analytics platform must encompass technologies from all three aspects, that is, ETL, text analytics algorithms, and visualization and user interaction. Each of these areas may contain their own specific technologies as discussed earlier.

Information Analytics Applications

To demonstrate how such analytics technologies can be used to address real-world problems, we describe a specific application that aims at deriving corporate brand and reputation insights by mining CGM content. Clearly, brand images and reputation are paramount to corporations, especially consumer-facing companies. It is extremely easy for a brand to become tarnished or become negatively associated with a social, environmental, or industry issue. This is true especially with the emergence of new forms of media, such as blogs, web logs, message boards, and web sites.

In a survey of the attitudes toward blogs, 77 percent of the respondents thought the regularly updated journals were a useful way to get insights into the products and services they should buy.[38] In a 2006 survey, 85 percent of respondents said word-of-mouth communication is credible, compared with 70 percent for public relations and advertising.[39] The new media allows consumers to spread information freely and at the speed of thought. By the time publicity has reached the press, it can be too late to protect the brand—only damage control is possible. Clearly, new analytical methods that leverage CGM content for early warnings on brand and reputation issues are needed. We present such an approach below, which uses four analytics components:

1. Broad keyword-based queries
2. Snippets
3. Annotation and taxonomy generation
4. Orthogonal filtering

The broad queries are similar to the "explore" queries described earlier. They are used to extract relevant information from the identified data sources. For example, if one were to study the consumer perceptions about a set of chocolate brands, one can issue a general keyword query such as the brand name to pull all relevant data from all data sources (e.g., blog feeds, news feeds, Web, and internal call center databases), similar to issuing Google search queries. Such broad queries are used to capture sufficient information about the target entities to be analyzed, such as brands and corporations. But they need not be extremely precise. Subsequent analysis steps will further process and filter the data. Many alerting systems are using such keyword search–based technologies alone to identify alerts. They often are ineffective due to the sheer volume of the returned data set.

Once the content is acquired, the second level of content filtering is called text *snippetization*. This is an important technique for analyzing Web content, since Web contents often are noisy. They may cover diverse topics in one document, even though only a few sentences might be relevant to the analysis subject. A snippet is a small text segment around a specified keyword. The text segment can be defined by sentence boundaries, or the number of words. In general, snippets are built around core keywords (e.g., brand names or corporation names). Snippets also reduce the total volume of data that users must read by focusing on the text segments that are relevant to the topic, rather than the whole documents at all times. In our evaluations, we found that snippetization can reduce the text size to be read by more than half.

After snippetization, taxonomy generation and annotation technologies are used to extract and identify brands and issues/topics from snippets. Brand annotators extract brand names from text snippets, and hot-word annotators extract hot issues about the brands from the snippets. Such two types of annotators can be used to form two new taxonomies on the data, that is, brands and hot issues. Finally, "orthogonal filtering" technique[40] is applied to identify interesting alerts with a high degree of accuracy by joining the above two constructed taxonomies.[41] Compared to typical corporate brand alert systems, which are based on keyword search technologies alone, such analytics-driven approach reduces massive amount of information down to a handful of alerts. This is especially true for many common brands and corporate names (e.g., from thousands of articles a day down to a dozen a day).

Case Study and Evaluation

The following case study shows the effectiveness of the analytics capabilities in a real-world situation. In this case study, we show how a company used such solutions to quickly detect significant consumer blog buzz after their product launch. In addition, the company can analyze where the buzz was coming from and how the buzz was evolving over the Internet. Such monitoring and analysis resulted in the change of company actions regarding a product launch in about one week. Because of this, the company's reputation and brand image ultimately improved significantly over the year, since other competitors could not take actions in a timely manner when faced with similar situations. Exhibit 14.3 shows the overall timeline of the event in details. We also list the detailed sequence of actions and events in seven days for this particular case:

- *Day 1*. The company announced products with specific ingredients that are unfriendly to certain ethnic communities.

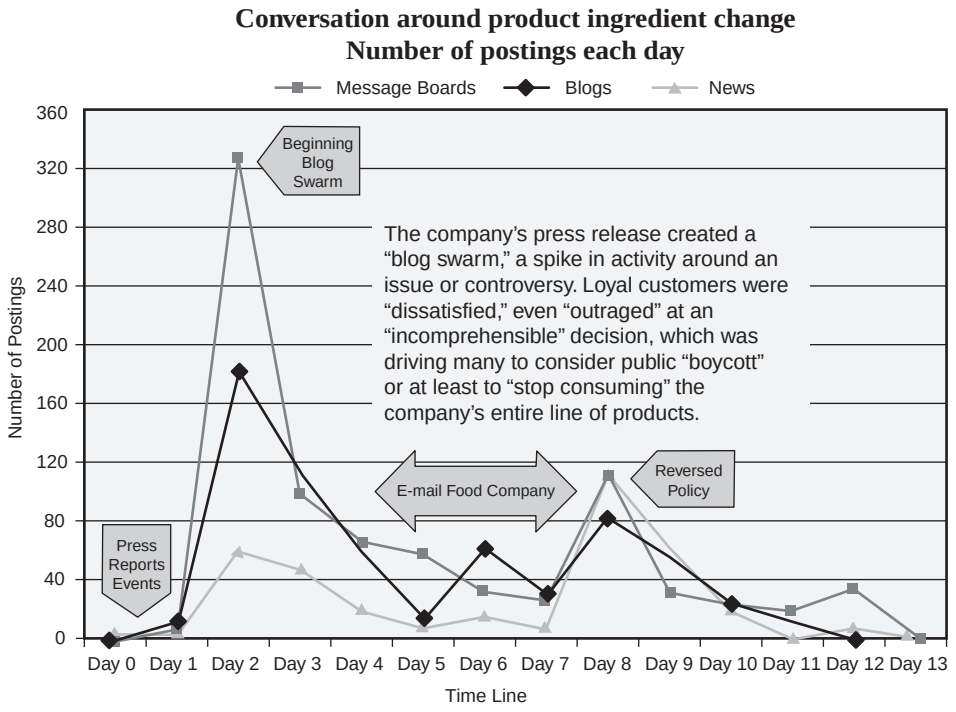


EXHIBIT 14.3 Contagion Effect Is Forcing Companies to Defend Themselves against "Consumer" Media Blog Swarms

- *Days 1 and 2.* Blogswarm (a spike of in activity around an issue or controversy) of protest were observed. The company analysts identified where the buzz came from and determined that many ethnic communities were outraged by the product announcements.
- *Days 3 to 6.* The ethnic communities sent e-mails and phone calls threatening to boycott company's products and stop consuming them all together.
- *Day 7.* Company reversed decision and apologized to the community.

FUTURE ANALYTICS TECHNOLOGIES

Besides the analytics technologies and their applications in real-world problem domains, we see that the future generations of analytics will need to deal with specific environmental forces. The following is a summary of such environmental forces and the associated analytics challenges:

- The massive growth of the information in all forms will create new challenges in the scalability and performance of analytics technologies. This is especially true when exponentially growing Web content is leveraged. Today's analytics technologies may need to be redesigned to run on massively parallel infrastructures to enable high speed and high scalability.

- The ecosystem complexity and dynamics require companies to understand many entities in their stakeholder network clearly before making decisions. Such social network analysis will become critical to all enterprises. We believe that social network analysis will become one of the most significant analytics technologies for enterprises. Today's social network analysis is still in its infancy. Many issues such as scale of the network and the dynamics of the networks render existing algorithms to fail.
- The development of new information platforms such as the Web and Internet makes it possible for everyone to benefit from analytics. We believe that analytics will become inherent and ubiquitous. To reach such a state, new technologies are needed to make analytics accessible, consumable, and attractive to users. To this end, we believe that new visualization technologies will be needed.

CONCLUSION

This chapter analyzed the information trends over the past 40 years and the associated technologies. This analysis indicates that our current and future generations will require analytics to be part of everyone's life. We provided an overview of the existing analytics technologies in the data-mining and text-mining space. We also showed how such analytics can be applied to address real-world problems through a case study. We outlined the future challenges of the analytics technology, in the area of scalability, social network analysis, and visualization. In the future, we will exploit many of such areas in our research as well.

NOTES

1. D. Tapscott and D. Ticoll, *The Naked Corporation: How the Age of Transparency Will Revolutionize Business* (New York: Free Press, 2000).
2. Ibid.
3. T. H. Davenport, and J. G. Harris, *Competing on Analytics: The New Science of Winning* (Cambridge, MA: Harvard Business School Press, 2003).
4. R. Srikant, Q. Vu, and R. Agrawal, "Mining Association Rules with Item Constraints." *Proceedings of the 3rd Int'l Conf. on Knowledge Discovery in Databases and Data Mining.* Newport Beach, CA, 1997.
5. W. Frawley, G. Piatetsky-Shapiro, and C. Matheus, "Knowledge Discovery in Databases: An Overview." *AI Magazine,* (Fall 1992): 213–228.
6. P. Tan, M. Steinbach, and V. Kumar, *Introduction to Data Mining* (Boston: Addison-Wesley, 2005).
7. R. Agrawal, "Data Mining: Crossing the Chasm." Keynote at the *5th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining.* San Diego, CA, 1999.
8. R. J. Bayardo and R. Agrawal, "Mining the Most Interesting Rules." In *Proceedings of the 5th CAN SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining,* 1999.
9. D. Grossman and O. Frieder, *Information Retrieval: Algorithms and Heuristics,* 2nd ed. (New York: Springer, 2006).
10. R. Baeza-Yates, and B. Ribeiro-Beto, *Modern Information Retrieval* (Boston: Addison-Wesley Publishing, 1999).
11. C. D. Manning and H. Schutze, *Foundations of Statistical Natural Language Processing* (Cambridge, MA: The MIT Press, 1999).

12. P. Jackson and I. Moulinier, *Natural Language Processing for Online Applications: Text Retrieval, Extraction, and Categorization* (Amsterdam: John Benjamins Publishing, 2002).
13. D. Modha and S. Spangler, "Feature Weighting in K-Means Clustering." *Machine Learning* 52(3) (2003): 217–237.
14. S. Spangler, J. Kreulen, and J. Lesser, "Generating and Browsing Multiple Taxonomies over a Document Collection." *Journal of Management Information Systems* 19(4) (2003): 191–212.
15. IBM Ascential, http://ibm.ascential.com.
16. IBM DB2, Data Warehouse Edition. www-306.ibm.com/software/data/db2/dwe.
17. Kalido, Enterprise Data Warehousing. www.kalido.com.
18. Han and M. Kamber, *Data Mining: Concept and Techniques*. Morgan Kaufmann, 2000.
19. P. Domingos and M. Pazzani, "On the Optimality of the Simple Bayesian Classifier under Zero-One Loss." *Machine Learning* 29 (1997): 103–137.
20. Christopher J. C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition." *Data Mining and Knowledge Discovery* 2 (1998): 121–167.
21. Alan Agresti, *Categorical Data Analysis* (New York: Wiley-Interscience, 2002).
22. V. S. Y. Lo, "The True Lift Model." *ACM SIGKDD Explorations Newsletter* 4(2) (2002): 78–86.
23. A. Behal, Y. Chen, C. Kieliszewski, et al., "Business Insights Workbench—An Interactive Insights Discovery Solution." In *Proceedings of the 12th International Conference on Human-Computer Interaction,* 2007.
24. W. S. Spangler and J. T. Kreulen, *Minding the Talk* (Armonk, NY: IBM Press, 2007).
25. T. Gotz, and O. Suhre, "Design and Implementation of the UIMA Common Analysis System." *IBM System Journal* 43(3) (2004).
26. B. He, R. Wang, Y. Chen, A. Lelescu, and J. Rhodes, "BIwTL: A Business Information Warehouse Toolkit and Language for Warehousing Simplification and Automation. *Proceedings of the ACM SIGMOD*, Beijing, China, 2007.
27. S. Arya, D. M. Mount, N. S. Netanyahu, R. Silverman, and A. Y. Wu, "An Optimal Algorithm for Approximate Nearest Neighbor Searching in Fixed Dimensions." *Journal of the ACM* 45(6) (1998): 891–923.
28. J. Lafferty, A. McCallum, and F. Pereira, "Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data. In *Proceedings of 18th International Conference on Machine Learning* (San Francisco: Morgan Kaufmann, 2001), 282.
29. T. R. Leek, "Information Extraction using Hidden Markov Models," Master's Thesis, UC San Diego, 1997.
30. See T. Gotz and O. Suhre, "Design and Implementation of the UIMA Common Analysis System." *IBM System Journal* 43(3): 2004.
31. D. Modha and S. Spangler, "Feature Weighting in K-Means Clustering." *Machine Learning* 52(3) (2003): 217–237.
32. W. S. Spangler, J. T. Kreulen, and J. F. Newswanger, "Machines in the Conversation: Detecting Themes and Trends in Information Communication Streams." *IBM Systems Journal* (2006).
33. Fabrizio Sebastiani, "Machine Learning in Automated Text Categorization." *ACM Computing Surveys* 34(1) (2002): 1–47.
34. Press, W. et. al., *Numerical Recipes in C*. 2nd ed. New York: Cambridge University Press (1992): 620–623.
35. S. Spangler and J. Kreulen, "Interactive Methods for Taxonomy Editing and Validation." *ACM CIKM* (2002).
36. J. Kreulen, W. S. Spangler, and J. Lesser, "MindMap: Utilizing Multiple Taxonomies and Visualization to Understand a Document Collection." *HCCI* (2002).

37. For examples of such visualization techniques, see Behal et al., 2007.
38. BBC Report, www.cymfony.com/know_center_blog.asp, 2007.
39. Harris Interactive Inc., "Word of Mouth Marketing—A Strategy." www.eoecho.com/gregmagnus/2006/03/word-of-mouth-marketing/, 2006.
40. See note 38.
41. S. Spangler, Y. Chen, L. Proctor, et al., "COBRA—Mining Web for Corporate Brand and Reputation Analysis." In *Proceedings of Web Intelligence Conference,* 2007.

CHAPTER 15

Embedded Predictive Analytics: Transforming Risk Management from Review Function to Competitive Advantage

Jill Eicher

INTRODUCTION

Predictive analytics are technology-enabled analytic methods for determining which course of action will drive success. These methods enable a business to scientifically learn from its experiential data and then immediately apply that knowledge to current decision making and execution. While predictive analytics have broad application, this chapter will examine how they apply to risk management in the financial services industry.

EXECUTION RISK IN THE FINANCIAL SERVICES INDUSTRY

As the interdependencies of risk in the global financial markets evolve and mutate, making risk/reward decisions has become significantly more challenging for investors. In the financial services industry, risk is both friend and foe. Risk creates investment opportunities; it can also eliminate their rewards. While a decision to invest is based on an evaluation of the risk associated with an investment opportunity, the analysis typically does not factor in the risk that the decision may not be carried out. The value of the investment opportunity, however, is lost entirely if the decision is not executed.

The failure to execute an investment decision is an uncompensated risk, one that is increasingly worrying investors and regulators alike. Whether due to operational failure or counterparty default, execution risk imperils investment returns, reputations, profitability, and market stability. The conventional approach to managing execution risk is through root cause analysis based on loss event histories. Unfortunately, this review function approach to execution risk has done more to increase concerns about execution risk than to proactively manage it.

Predictive analytics provide the ability to detect and deter execution risk in real time. Technologies such as complex event processing, data stream management, and business activity monitoring enable linear, rules-based, and machine learning analytic methods to be embedded into the business processes of financial services organizations. Once embedded, predictive analytics serve as real-time sentinels preventing operational errors and counterparty delinquencies from incurring unnecessary costs and triggering loss events.

BUSINESS PROCESSES

The business processes of a financial services organization are best thought of as its DNA, the source code of its competitive edge. The effectiveness of business processes determines the quality and consistency of execution service delivery to clients. Measuring business process effectiveness provides a quantitative framework to understand the factors driving profitability and sustainability.

While the concept of measuring and analyzing business processes is relatively new in the service industry, the manufacturing sector has long been proficient in mining every last cent of value from business processes in the face of narrowing margins and global competition. Forward-thinking financial services organizations, however, are realizing that the data extracted to understand business processes can be used not only to root out inefficiencies, but also to proactively manage risk.

The counterparty selection process provides a good example of the value derived from analyzing business processes. Traders can often choose among several counterparties offering a desired security or contract with similar investment characteristics. Typically, the trader then takes several additional factors into consideration, such as:

- Financial strength based on stored credit rating data, regulatory financial filings, and earnings information.
- Current outstanding commission obligations based on in-house portfolio manager/analyst and client-directed targets.

While a review of static financial information and outstanding commission obligations have long sufficed, traders without more relevant and timely insight are at a disadvantage in today's environment of declining investment returns and increased market volatility. To make a good counterparty selection decision, the trader needs to know which counterparty is most likely to provide the requisite operational expertise as well as which one has the current financial wherewithal to meet contractual obligations.

Operational expertise can be measured in terms of on-time settlement and profitability. By digging into the firm's transaction history, data about the execution history of each counterparty can be analyzed transaction by transaction in great granularity. A statistical analysis of the settlement process reveals which counterparties are operationally effective in terms of high on-time settlement frequency, as well as which ones do not incur unnecessary rework, remediation, or extended financing costs.

Streaming news, market data feeds, and electronic regulatory sources supply information about the current financial strength of counterparties. Real-time monitoring of financial sources for default events and capitalization changes provide an up-to-the-minute perspective on the financial condition of counterparties.

The ability to select a counterparty based on operational expertise and current financial strength increases the probability of effective execution. By knowing which counterparty is most likely to execute a transaction effectively, the trader can increase efficiency and decrease exposure to execution risk. This is a competitive edge for a financial services firm, particularly in a changing landscape of counterparties.

It is not possible, however, for a trader to manually comb through transaction histories and cull minute-to-minute financial information for each counterparty selection decision. Predictive analytics provide the ability to perform these analyses routinely and efficiently. These analyses provide the trader with better information to make a better decision in selecting a counterparty.

The ability to optimize business processes by applying inferences derived from the business's operating infrastructure is a distinguishing characteristic of predictive analytics. What's more, these algorithmic-driven analytics are designed to continuously learn from inference by adapting the underlying analytic method to the derived inferences. This means the more data crunched, the better the analysis. By embedding predictive analytics into business processes, a business taps into a proprietary source of competitive information.

Predictive analytics present an enormous opportunity for the development of risk management practices to go beyond reviewing operational and counterparty exposures to quantitatively managing execution risk in real time. While predictive analytics have been used for many years in the financial services industry to fuel algorithmic trading programs, recognition of their transformative potential for risk management is just emerging.

PREDICTIVE ANALYTICS: TECHNOLOGY-ENABLED ANALYTIC METHODS

The ability to embed predictive analytics into the business processes of financial services firms is changing the competitive landscape for investors. Enabling technologies have reset the analytic time zone; increased computing power in terms of access, speed, and volume; and commoditized data storage. Data-mining techniques have tapped into new domains of source data enriching predictive models and amplifying analytic methods. Collectively, these advances are creating a dynamic analytic process of continuous learning and improvement in making and executing investment decisions (see Exhibit 15.1).

As a result, utilization of predictive analytics is expanding as investors discover that the difference between profit and loss is no longer determined solely by the decision of what to buy or sell, but also by how effectively the decision is executed. The forward-looking perspective and actionable nature of predictive analytics allows financial services firms to extend competitive advantage through the execution of the investment decision. The potential of these technology-enabled analytic methods have particular import for risk management methodologies previously relegated to a retrospective review function based on subjective assessment practices and historical data.

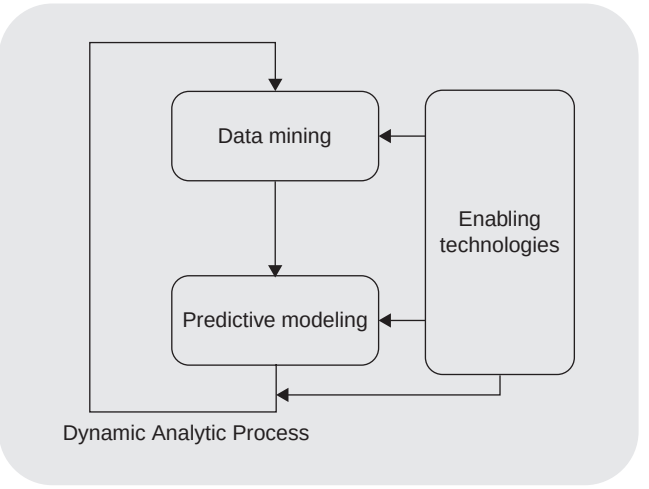


EXHIBIT 15.1 Predictive Analytics

Enabling Technologies

Advances in technology architecture design and computing power and data storage feasibility in processing technologies have converged to redefine analytic study. Innovations in service-oriented and event-driven technology architectures powered the advent of processing technologies such as complex event processing, business activity monitoring, and data stream management. Technology architectures and processing technologies evolved to handle unparalleled volumes of disparate data in lightning speed. These innovations changed the analytic time horizon, making it possible to examine not only "what happened," but also "what is happening now."

Innovators developed processing models and programming language techniques that could comb vast volumes of live data in milliseconds and identify complex sequences of events with temporal parameters. These advances introduce the potential to eliminate risk management's reliance on database-driven analysis and corresponding retrospective focus. Broadening the time horizon and expanding the experiential data set available for predictive analysis, paved the way for technology-enabled analytic methods to exploit both historical data and current events to forecast "what is most likely to happen in the future."

Technology Architectures Service-oriented and event-driven architectures increased the efficiency of the business services and systems with a business infrastructure. They also expanded business querying accessibility from databases and data warehouses to all business infrastructure systems and services. The sophistication of these technology architectures has made it possible to understand how well the business processes and operating infrastructure of financial services firms are performing.

- *Service-oriented architecture (SOA).* A technology infrastructure that allows business services to work together and/or communicate with one another. Within this infrastructure, services operate more like business functions and processes by

sharing data while in operation. In addition to speeding the delivery of services to traders, analysts, risk managers and operations teams, SOAs make information about their operation and use available for analysis. Of particular import to risk analysis is the access SOAs provide to transactional activity histories and infrastructure audit logs. This access permits analysis of how well the business processes of the firm are performing.

Event-driven architectures (EDAs). Designed to allow businesses to monitor, analyze, and act on events impacting their operation. An event is an occurrence that impacts business strategy, for example, a credit rating change or the sale of a security. Event-driven architectures funnel insights derived from current events into models for further analytic study.

Processing Technologies Complex event processing, business activity monitoring, and data stream management processing technologies allow investors to manage dynamics impacting the business in real time. This facilitates early detection and prompt response to operational issues compromising profitability and execution.

Complex event processing (CEP). A sophisticated event-tracking and pattern analysis technology that facilitates the management and analysis of high-volume, real-time business activity. It does so by using sophisticated programming languages to detect complex events within a context of specific pattern constraints. In addition to the detection of event sequences, CEP technology analyzes events and triggers action when proscribed. An event can be thought of as an occurrence happening either externally to a firm or within a firm's technology infrastructure. A trader executing an order to buy 100,000 shares of Google at $324.50 is an example of an event. The sequence of (1) news that Microsoft will not buy Yahoo!; (2) Yahoo! stock price drops to 76.80; and (3) the trader buys 100,000 shares of Google at $324.50 is an example of a complex event. CEP technology allows financial services firms to assess the implications of an event, determine the optimal action to be taken, and execute that action, all within milliseconds. Moreover, CEP technology is enabling investors to act on events without human intervention.

Business activity monitoring (BAM). A real-time activity measurement technology that monitors and evaluates the performance of a business infrastructure. BAM technology manages event data compiled from the operating infrastructure of a firm and provides status and alert information about business process service delivery. Business activity can be either an individual business process or a sequence of activities stemming from various applications and systems. Trade confirmation is an example of a business activity. Business activity monitoring provides real-time performance summaries of a firm's business operations. By alerting investors to changes in operational status or counterparty performance, action can be taken to prevent operational failure or counterparty default.

Data stream management (DSM). A continuous querying technology that facilitates the management of real-time, online data streams and the deployment of continuous queries on them. To perform online analysis of arriving data

in real-time, DSM technology creates temporal data models and then applies complex filtering and query semantics to evaluate each data item in a continuous data stream. A data stream is composed of a continuous, real-time sequence of items. A market data feed is an example of a data stream. DSM allows financial services businesses to continuously access and process large volumes of real-time and historical data to test insights, theories, and scenarios using from analytic queries.

Data Mining

Efficient access to previously untapped source data has unlocked new domains of information discovery to advance analytic study. The ability to systematically analyze raw, unobserved, and unformatted data spurred investigation into the relationships, trends, and patterns between new data sets as well as traditional source data. These advancements made it possible to use computer-driven methods to evaluate what is and what is not working in business strategies and operations, as well as what will work in the future given historical and current precedents.

By developing sophisticated algorithms that dissect both structured and unstructured data, data-mining developers expanded the scope of source data for predictive analysis to extend beyond traditional database and data warehouse sources. These new data-mining techniques introduced real-time source data into the analytic process. With the ability to extract useful information from both current and historical data, data mining evolved into a continuous information discovery process. Data mining introduces financial services firms to a proprietary source of strategic information to fuel the statistical analysis of optimal risk/reward decision making and execution.

Source Data Data-mining techniques have long focused on the analysis of structured data, typically data organized, formatted, and stored in a relational database. The process of structuring the data imposes a predetermined order to the data and its known relationships. The order is designed to facilitate anticipated querying and information retrieval needs. It also provides a context for the data that serve as a permanent association. The static nature of the ordering process has generally limited the usefulness of the information derived from structured data.

Metadata, however, is a new form of structured data that has introduced a valuable source data for predictive analysis. Previously known and used predominantly by IT experts, metadata is data assigned to other data elements to describe the data. It can be thought of as data about the data. A transaction identification number and time stamps assigned to securities transactions are examples of metadata. Metadata is particularly useful in risk analysis because the descriptors assigned to the data enrich the source data for analysis without biasing the analysis.

The need to understand current events and operational performance in the context of business strategy effectiveness is driving the development of data-mining techniques to be applied to semistructured and unstructured data. Unlike structured data, the order of semistructured data is not imposed by an explicit data model. Instead, its order is set locally, providing some degree of implicit structure yet not imposing the structure unilaterally. Stock tick data and a series of related spreadsheets are

examples of semistructured data. Unstructured data, however, exists without any kind of definition. A number is simply a number, and a word is simply a word. A spreadsheet and an e-mail are examples of unstructured data. In preparation for analysis, unstructured data needs some degree of human intervention to be made computer ready.

The introduction of new source data domains allows financial services firms to learn from more of their own operating data. It also offers the potential to end the reliance on generic industry and loss event data for risk analysis. The ability to perform granular analysis on experiential data provides financial services firms with the ability to have better information to better manage their businesses.

Information Discovery The volume and complexity of data being processed by financial services firms exceeds the human capacity to analyze. The objective of data mining is to use computer-driven techniques to make the analysis of large datasets possible. These techniques are designed to gather insights and inferences about patterns, relationships, and correlations from data and translate them into useful information. In translating the extracted data, new information is discovered.

There are many different approaches to the information discovery process. Some of the most common are quantitative, classification, visualization, and machine learning. They can be thought of broadly as intelligent learning techniques used to discover information from datasets about the determinants of success, including optimal behavior and results.

- *Quantitative*. Probability and statistics are the two primary quantitative analysis approaches. Probability techniques compare different information scenarios and assign a probability to each outcome. Statistical techniques generalize patterns in datasets and develop rules from the patterns.
- *Classification*. There are many classification approaches including the most widely used Bayesian, decision-tree analysis, and pattern recognition. Classification techniques group data according to similarities or categories.
- *Visualization*. The use of graphical tools to interpret and illustrate complex relationships in multidimensional data.
- *Machine learning*. The application of induction algorithms to learn from experience and to adapt automatically when new factors are introduced that change expected future success (i.e., performance and profitability).

The ability to apply intelligent learning techniques to large and complex datasets uncovered information never imagined. It was not long until these techniques were harnessed to identify the determinants of success. The evolution of the information discovery process also makes it possible to identify the determinants of execution risk. For example, a counterparty's error and settlement rate history might help predict the likelihood of on-time settlement.

Data mining is about gathering information on determinants that influence future successful outcomes. The next section on predictive models is about applying that information to optimize successful outcomes.

Predictive Modeling

New proficiencies in developing predictive models coupled with innovations in technology-enabled analytic methods are revolutionizing statistical analytic study. The ability to algorithmically build predictive models using determinative variables extracted from experiential data led to programming computers to learn mechanically from the analyses they perform. These advances in predictive modeling introduced computer-driven statistical reasoning.

The functionality of predictive models was extended well beyond calculating probabilities when advanced algorithms were developed to turn inference into learning and action. This was achieved by applying technology advances in architecture, processing, source data extraction, and information discovery to analytic methods. Thanks to the evolution of technology-enabled analytic methods in predictive modeling, statistical analysis is no longer limited to confirming hypotheses, but now proactively informs. The ability to interpret data in real time and automatically generate optimal next steps provides financial services firms with the ability to proactively manage risk.

Model Development A predictive model is a set of algorithms that turns inferences, statistically derived from experiential data, into action in order to advance business strategy and execution. The algorithms are instructions in the form of mathematical formulas that specify: (1) what problem the predictive model is solving; (2) how to analyze the experiential data; (3) how to determine which actions would optimally solve the problem; and (4) how to apply and adapt the inferences extracted from the analytic process to continuously hone the predictive model. Collectively, the algorithms driving predictive models are based on the determinants of future behavior or results identified by data-mining algorithms.

Creating a predictive model is an iterative process. Determinants and inferences from experiential data drive the initial development of a predictive model as well as its ongoing refinement. More data increases the model's precision.

The utilization of determinants and inferences derived from both historical and current experiential data on a continuous basis distinguishes predictive models from other analytic frameworks. This feature is at the core of the predictive model's functionality and provides the foundation for how it generates organic and relevant analysis. Predictive modeling eliminates the problems of forensic querying, stale data, and obsolete analytic frameworks associated with traditional statistical analysis and instead make computer-driven statistical reasoning possible.

Analytic Method Analytic methods are employed in predictive models to generate information that can be used to drive successful outcomes. They do this by examining how specific business dynamics relate to past, present, and future activity. Unlike in data mining, where analytic methods typically classify all correlations found in the data, in predictive models, analytic methods are used to search for specific causal relationships and determinants. The analytic methods used in predictive models can broadly be grouped into three categories: linear, rules-based, and machine learning.

1. *Linear*. Regression techniques are the predominant analytic method of predictive analysis. They are used to create mathematical equations to model the relationships between dependent and independent variables. Linear analytic

methods produce predictive equations designed to measure the predictive capability of independent variables.

- *Linear regression*. This method analyzes the relationship between predictor variables (i.e., the independent variables that influence an outcome, represented by the dependent variable). In the equation, the variables are used to express the relationship as a linear function. Linear regression equations solve explicitly.
- *Logistic regression*. Using a predictive equation, logistic regression is an iterative regression method that is used when the outcome variable is indicative (as opposed to a quantitative variable, which characterizes linear regression). A sequence of trial equations continues until fit is achieved.

2. *Rules-based*. An "if-then" method based on two components, a condition, and an action. The condition is often used to identify characteristics of a data set that may serve as predictors. Rules-based predictors function similarly to an independent variable in multivariate statistics. A security description is an example of a characteristic of a security transaction and a security description error is an example of a predictor. The probability of the security transaction settling is an example of an action, which specifies the outcome.
   - *Decision trees*. The most widely used rules-based analytic method, decision trees are utilized for classification and pattern recognition. They are particularly effective when there is a large field of variables to understand. Decision trees divide large data sets into successively smaller data sets by applying a sequence of simple decision rules.
3. *Machine learning*. These methods combine sophisticated computer science and statistical techniques to produce learning algorithms that automatically use the data being analyzed to enhance the analytic method.
   - *Neural networks*. A highly advanced, nonlinear statistical technique used to model complex data sets, particularly when the relationship between the data set and the outcome is unknown. A distinguishing feature of a neural network is the ability to learn from the data in a way similar to human cognition. Neural networks have a facile ability to derive meaning from complicated or imprecise data.
   - *Memory-based reasoning*. A sophisticated similarity-based technique used to answer questions or solve problems by employing analogous variables. Drawing from an ability to analyze raw data sets, memory-based reasoning determines which set of variables most closely resemble current criteria. Often associated with nearest neighbor approaches, this method relies more on iterative analysis than a strong domain model, inference, or rules. Memory-based reasoning is distinguished by using additions to the data set to learn and adapt the method.
   - *Support vector machines*. A supervised learning method used to find complex patterns and transform them into organized data sets by applying sophisticated classification and regression methods. They are best known for their ability to apply linear classification techniques to nonlinear classification problems.
   - *Naïve Bayes*. A special form of Bayesian probability method used primarily for classification and clustering, Naïve Bayes is particularly useful in large or real-time data set applications. Naïve Bayes is differentiated by its prompt generation of inference. It is most often utilized in predictive models driven by a numerous independent variables.

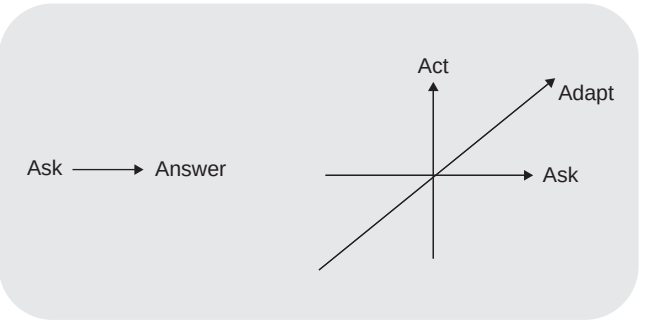


EXHIBIT 15.2 Traditional Statistical Analytic Process vs. Dynamic Analytic Process—First Example

- *Genetic algorithms.* Mimicking a Darwinian survival-of-the-fittest evolutionary approach to find the best, whether applied to patterns, relationships, or correlations, these algorithms select the best and eliminate the worst, generating mutations of the best to create even better algorithms. Genetic algorithms use selection, recombination, and mutation to "breed" a solution to a problem. They are particularly useful in finding optimal parameters in complex data sets.

Dynamic Analytic Process

Thanks to advances in enabling technologies, data mining, and predictive modeling techniques, predictive analysis is a dynamic analytic process. A continuous stream of new data is pumped into the predictive analytic engine to identify the most current determinants of success. From this process, the optimal course of action is discovered. In predictive analysis, the work is never done because there are always new data to analyze. Predictive analysis, therefore, can be thought of as computer-driven continuous learning.

Exhibits 15.2 and 15.3 show the contrast between the linear "ask" and "answer" process of traditional statistical analysis, and the predictive analysis using a multidimensional process of "ask," "act," and "adapt." For this reason, predictive analysis is particularly well suited to the analysis of the temporal networks of risk that characterize today's global financial markets. Financial services firms are just beginning to exploit the proprietary information advantages derived from predictive analysis for risk/reward decision making and execution.

The dynamic analytic process, fueled by predictive analytics, provides financial services firms with a sustainable information advantage to navigate constantly changing business dynamics and financial markets.

CONCLUSION: MANAGING RISK COMPETITIVELY

In an era of diminishing returns and more volatile global financial markets, investors can no longer afford the uncompensated risk of operational failure and counterparty default. Likewise, these uncompensated execution risks increasingly tax the bottom

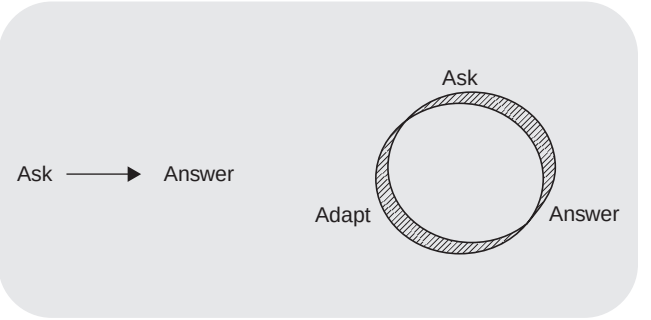


EXHIBIT 15.3 Traditional Statistical Analytic Process vs. Dynamic Analytic Process—Second Example

lines of financial services firms in the form of rework, remediation, and extended financing costs as volume and complexity soar.

Enterprising managers can reduce uncompensated execution risks by embedding predictive analytics into their business infrastructures. Once embedded, these technology-enabled analytic methods serve as real-time sentinels preventing operational errors and counterparty delinquencies from incurring unnecessary costs or triggering loss events.

The ability to proactively manage execution risk enhances investment returns for investors and increases profits to shareholders. In contrast to the traditional retrospective approach to risk management, this capability enables financial services firms to extend investors' competitive advantage through the execution of the investment decision.

CHAPTER 16

Reducing the Financial Risks in Litigation and Legal Discovery

Anthony Tarantino, Ph.D.

BACKGROUND

Over the past year, I have made presentations in Europe and the United States to very diversified audiences arguing that the United States can now be considered the most litigious society in history. No one has ever challenged this assertion, and to the contrary attendees have provided both U.S. and European examples to embellish the point. A few interesting statistics and factoids will help set the stage for this discussion based on research of U.S. companies by Gartner Research,[1] the law firm of Fulbright and Jaworski, LLP,[2] and my own research:

- U.S. companies with $1 billion plus revenues are involved in over 500 cases, with an average of 50 new disputes emerging each year.
- The typical cost is $1.2 to 1.4 million per suit, before any judgments or settlements.
- About 70 percent of companies have initiated their own legal actions.
- Nearly 40 percent had at least one suit of over $20 million launched against them last year.
- Legal discovery represents 70 percent of litigation costs in most major cases.
- About 50 percent of all the world's lawyers are in the United States.
- There are over one million lawyers in the U.S.
- U.S. plaintiffs filed 30 million new lawsuits last year or 82,000 per day.
- Labor law disputes are the leading cause of suits followed by contract disputes.

The poisonous litigious environment in the United States has become a major fear factor among its trading partners, even those that enjoy very high legal protections and civil rights. The problem is compounded by U.S. court rulings that hold foreign companies to U.S. legal standards as long as they enjoy the benefits of doing business in the United States. European Union (EU) courts have backed these U.S. court decisions, which have even trumped EU privacy protections.

Legal discovery is the process of finding information that was not previously known. It is compulsory to share this information, upon the request to the other parties in a litigation.

Electronic discovery refers to the discovery of electronic records, documents, and metadata. Electronic documents include e-mail, Web pages, word processing files, computer databases, and virtually anything that is stored on a computer, along with their reference metadata. Technically, documents and data are electronic if they exist in a medium that can be read only through the use of computers. Such media include cache memory, magnetic discs (such as computer hard drives or floppy discs), optical discs (such as DVDs or CDs), and magnetic tapes.

Electronic discovery is often distinguished from paper discovery, which refers to the discovery of writings on paper that can be read without the aid of some devices. Is digital information different? Computer files, including e-mails, are discoverable. However, courts are not persuaded by the plaintiffs' attempt to equate traditional paper-based discovery with the discovery of e-mail files. There are important differences between the two. Chief among these differences is the sheer volume of electronic information. E-mails have replaced other forms of communication besides just paper-based communication. Many informal messages that were previously relayed by telephone or at the water cooler are now sent via e-mail and now instant messages. Many users of electronic communications now consider e-mail the new snail mail, the term reserved for paper mail in the past.

THE SEDONA CONFERENCE AND THE NEW RULES OF CIVIL PROCEDURE

The Sedona Conference 14 Principles

The Sedona Conference® is a nonprofit, research, and educational institute dedicated to the advanced study of law and policy in the areas of antitrust law, complex litigation, and intellectual property rights. The Sedona Conference Working Group combed the thoughts of in-house counsel, outside attorneys, and judges before settling on 14 principles for electronic discovery.[3] They are the foundation for the Federal Rules of Civil Procedure (FRCP) and go as follows:

1. Electronic data and documents are potentially discoverable. Organizations must properly preserve electronic data and documents that can reasonably be anticipated to be relevant to litigation.
2. When balancing the cost, burden, and need for electronic data and documents, courts and parties should apply a balancing standard embodied in federal codes and state laws equivalents, which require considering the technological feasibility and realistic costs of preserving, retrieving, producing, and reviewing electronic data, as well as the nature of the litigation and the amount in controversy.
3. Parties should confer early in discovery regarding the preservation and production of electronic data and documents when these matters are at issue in the litigation, and seek to agree on the scope of each party's rights and responsibilities.
4. Discovery requests should make as clear as possible what electronic documents and data are being asked for, while responses and objections to discovery should disclose the scope and limits of what is being produced.

5. The obligation to preserve electronic data and documents requires reasonable and good-faith efforts to retain information that may be relevant to pending or threatened litigation. However, it is unreasonable to expect parties to take every conceivable step to preserve all potentially relevant data.
6. Responding parties are best situated to evaluate the procedures, methodologies, and technologies appropriate for preserving and producing their own electronic data and documents.
7. The requesting party has the burden on a motion to compel to show that the responding party's steps to preserve and produce relevant electronic data and documents were inadequate.
8. The primary source of electronic data and documents for production should be active data and information purposely stored in a manner that anticipates future business use and permits efficient searching and retrieval. Resort to disaster recovery backup tapes and other sources of data and documents requires the requesting party to demonstrate need and relevance that outweigh the cost, burden, and disruption of retrieving and processing the data from such sources.
9. Absent a showing of special need and relevance, a responding party should not be required to preserve, review, or produce deleted, shadowed, fragmented, or residual data or documents.
10. A responding party should follow reasonable procedures to protect privileges and objections to production of electronic data and documents.
11. A responding party may satisfy its good-faith obligation to preserve and produce potentially responsive electronic data and documents by using electronic tools and processes, such as data sampling, searching or the use of selection criteria, to identify data most likely to contain responsive information.
12. Unless it is material to resolving the dispute, there is no obligation to preserve and produce metadata absent agreement of the parties or order of the court.
13. Absent a specific objection, agreement of the parties or order of the court, the reasonable costs of retrieving and reviewing electronic information for production should be borne by the responding party, unless the information sought is not reasonably available to the responding party in the ordinary course of business. If the data or formatting of the information sought is not reasonably available to the responding party in the ordinary course of business, then, absent special circumstances, the costs of retrieving and reviewing such electronic information should be shifted to the requesting party.
14. Sanctions, including spoliation findings, should only be considered by the court if, upon a showing of a clear duty to preserve, the court finds that there was an intentional or reckless failure to preserve and produce relevant electronic data and that there is a reasonable probability that the loss of the evidence has materially prejudiced the adverse party.

Legal Terms from the Sedona Conference Used in Legal Discovery

Understanding technical terms is necessary in mastering electronic evidence. The Sedona Conference also published a glossary of words and phrases used in electronic discovery.[4] Here are some examples:

- *Distributed data.* Distributed data is that information belonging to an organization which resides on portable media and nonlocal devices such as home computers, laptop computers, floppy discs, CD-ROMS, personal digital assistants (PDAs), wireless communication devices (e.g., Blackberry), zip drives, Internet repositories such as e-mail hosted by Internet service providers or portals, Web pages, and the like. Distributed data also includes data held by third parties such as application service providers and business partners.
- *Forensic copy.* A forensic copy is an exact bit-by-bit copy of the entire physical hard drive of a computer system, including slack and unallocated space.
- *Legacy data.* Legacy data is information the development of which an organization may have invested significant resources to and that has retained its importance, but has been created or stored by the use of software and/or hardware that has been rendered outmoded or obsolete.
- *Residual data.* Residual data (sometimes referred to as *ambient data*) refers to data that is not active on a computer system. Residual data includes (1) data found on media free space; (2) data found in the file slack space; and (3) data within files that have functionally been deleted in that it is not visible using the application with which the file was created, without use of undelete or special data recovery techniques.
- *Migrated data.* Migrated data is information that has been moved from one database or format to another, usually as a result of a change from one hardware or software technology to another.
- *System data,* or information generated and maintained by the computer itself. The computer records a variety of routine transactions and functions, including password access requests, the creation or deletion of files and directories, maintenance functions, and access to and from other computers, printers, or communication devices.
- *Backup data*, generally stored offline on tapes or disks. Backup data are created and maintained for short-term disaster recovery, not for retrieving particular files, databases, or programs. These tapes or discs must be restored to the system from which they were recorded, or to a similar hardware and software environment, before any data can be accessed.
- *Residual data* that exist in bits and pieces throughout a computer hard drive. Analogous to the data on crumpled newspapers used to pack shipping boxes, these data are also recoverable with expert intervention.
- *Active, online data.* Online storage is generally provided by magnetic disc. It is used in the very active stages of an electronic record's life—when it is being created or received and processed, as well as when the access frequency is high and the required speed of access is very fast (i.e., milliseconds). Examples of online data include hard drives.
- *Nearline data.* This typically consists of a robotic storage device, (robotic library) that houses removable media, uses robotic arms to access the media, and uses multiple read/write devices to store and retrieve records. Access speeds can range from as low as milliseconds if the media is already in a read device, up to 10 to 30 seconds for optical disc technology, and between 20 and 120 seconds for sequentially searched media, such as magnetic tape. Examples include optical discs.
- *Offline storage/archives.* This is removable optical disc or magnetic tape media, which can be labeled and stored in a shelf or rack. Offline storage of electronic

records is traditionally used for making disaster copies of records and also for records considered "archival" in that their likelihood of retrieval is minimal. Accessibility to offline media involves manual intervention and is much slower than online or near-line storage. Access speed may be minutes, hours, or even days, depending on the access effectiveness of the storage facility.

- *Metadata.* Metadata is information about a particular data set that describes how, when, and by whom it was collected, created, accessed, and modified and how it is formatted. Some metadata, such as file dates and sizes, can easily be seen by users; other metadata can be hidden or embedded and is unavailable to computer users who are not technically adept. Metadata is generally not reproduced in full form when a document is printed. (Metadata is typically referred to by the not highly informative shorthand phrase "data about data," describing the content, quality, condition, history, and other characteristics of the data.) Metadata supporting may be larger than the file itself. For example, 80 application and system metadata fields are tracked for MS Word document files. Metadata has become a critical component in legal discovery as litigants become more technically expert. Without metadata, it is often impossible to establish authenticity and relevancy of records and documents. To paraphrase the old newspaper adage: metadata provides litigants with the who, what, when, and where behind any given document. It does not provide the why. Metadata can be classified into two types:
  1. *Application metadata* is embedded within the file. It describes the file and moves with the file when it is copied.
  2. *System metadata* is an analogous to a library card catalog. It is stored and maintained external to the file.

Federal Rules of Civil Procedure: December 2006

The rules and methodology used in legal discovery were greatly clarified with the December 2006 approval by the Supreme Court and Congress of new FRCP.[5] The FRCP govern civil procedure in U.S. district (federal) courts, date back to 1938, and have been revised 10 times over the years. While U.S. states determine their own rules that apply in state courts, most states have adopted rules that are based on the FRCP.

Before the FRCP, *common-law pleading* was more formal, traditional, and demanding in its phrases and requirements. In contrast, the FRCP is based on a legal construction called *notice pleading*, which is less formal, created and modified by legal experts, and far less technical in requirements. In notice pleading, the same plaintiff bringing suit would not face dismissal for lack of the exact legal term, so long as the claim itself was legally actionable. The policy behind this change is to simply give "notice" of your grievances and leave the details for later in the case. This acts in the interest of equity by concentrating on the actual law and not the exact construction of pleas. Some states, such as California, use an intermediate system known as *code pleading*. Code pleading is an older system than notice pleading and is based on legislative statute. It tends to straddle the gulf between obsolete common-law pleading and modern notice pleading. Code pleading places additional burdens on a party to plead the "ultimate facts" of its case, laying out the party's entire case and the facts or allegations underlying it. Notice pleading, by contrast, simply requires a "short and plain statement" showing only that the pleader is entitled to relief (FRCP

8(a)(2)). One important exception to this rule is that when a party alleges fraud, that party must plead the facts of the alleged fraud with particularity (FRCP 9(b)).

A summary of the specific rules follows:

- *Rule 16(b)* now makes provisions to meet in advance of the trial to discuss discovery issues related to electronically stored information.
- *Rule 26(a)(1)* states that litigants must provide the names of holders of its relevant information and a copy or description of the data it will use to the other parties in the litigation, without awaiting a discovery request. This needs to be done in a timely manner, but the determination of timely is left to judges.
- *Rule 26(b)(2)(B)* deals with the issues of the discovery of information that is not reasonably accessible because of undue burden or cost. There are protections from cost prohibitive discovery such as requesting all e-mails that a company generates rather than those specific to a case. Litigants need not search or produce electronically stored information (initially) from sources that are not reasonably accessible because of undue burden or cost. Judges can mandate cost shifting and/or cost sharing in cases where the information is needed but considered unduly costly to produce. Litigants must identify, by category and type, the sources containing potentially responsive information that they are not producing. Identifying a source as not reasonably accessible does not relieve the litigant of its common-law or statutory duties to preserve evidence. Examples of inaccessible sources under "current" technology include: magnetic backup tapes, legacy data that is unintelligible, fragmented data after deletion, unplanned output from databases different from designed uses. Even inaccessible information must be produced if ordered for "good cause" or if access to the source is shown not to be sufficiently difficult because of "undue burden or cost."
- *Rule 26 (b)(5)(B)* states that privileged information is protected in what is called a "clawback" and safe harbor provision in which litigants must promptly return, sequester, or destroy it upon its discovery. Judges may impose time limits to this process. Courts will look at five factors in considering a clawback: the reasonableness of the precautions taken to prevent inadvertent disclosures, the time to rectify the error, the scope of the production, the extent of disclosure, and overriding issues of fairness.
- *Rule 26(f)* touches on a wide range of issues including discussing any issues relating to preserving discoverable information at the pretrial meetings. As soon as practicable, litigants must confer and come to a consensus as to what is in scope and out of scope in what has become a critical meeting to develop a discovery plan. This includes the identification, sources, and forms of production for ESI, whether the ESI is reasonably accessible, the burden and cost of retrieving and reviewing such information, and finally resolving issues relating to claims of privilege, including postproduction assertion of privilege or work-product protection. A discovery plan should include knowing which data is where, actions taken to preserve it, time and effort to get to it, how it can be searched and retrieved, what is privileged, what will not be searched, and in what format and media it can be provided.
- *Rule 33* is amended to make it clear that the option to produce business records includes electronically stored information.
- *Rule 34* adds "electronically stored information" as a category subject to production. Rule 34 (b) permits a requesting party to specify the form or forms in

which electronically stored information (ESI) is produced. The court has coined the term *electronically stored information* as a category of discoverable information. ESI includes unstructured data, such as e-mail and instant messages, and structured data, such as customer, supplier, and item masters. Absent agreement or court order, electronically stored information must be produced in form or forms "in which it is ordinarily maintained" or in a "reasonably useable" form. Material metadata may require "native format" production. No type of ESI is excluded from the discovery process and many judges have become technically expert in mastering the types of ESI.

- *Rule 37* is amended to address the problem of the destruction of records as a result of the routine, good-faith operation of an electronic information system. The rule is not intended "to provide a shield for the destruction of information related to a litigation." There is no penalty for purges as part of normal, routine, and good-faith operations, but once a suite is filed, litigants must stop the purge process or face sanctions. Rule 37 defines routine losses as "the ways in which such systems are generally designed, programmed, and implemented to meet the party's technical and business needs." A defensible routine would include the following: deletion repetitively occurs in a verifiable periodicity specified by an enforced policy; similar procedures followed for similar deletions events; scheduled and predictable, not "event driven," best if tied to a records management policy disposition schedule specified in a file plan.
- *Rule 45* is amended to provide for subpoenas regarding electronically stored information as well as paper documents. In specifying the form of production, Rule 45 acknowledges that electronic information can be sought through a subpoena as well as traditional discovery requests. The proposed amendments to Rule 45 incorporate changes to Rule 26 and 34 to provide parameters for production of electronic data through a subpoena, plus the ways in which such systems are generally designed, programmed, and implemented to meet the party's technical and business needs. In some cases, testing and sampling of electronic documents is used to determine the ultimate burden and costs.

U.S. COURT RULINGS UNDER THE NEW FRCP

There are several court rules that demonstrate the impact of the FRCP and current litigious environment. Many of these case law examples are now accepted as precedent setting[6,7,8]:

Discovery cost shifting. In two major cases, *Rowe Entertainment, Inc. v. William Morris Agency, Inc.* and *Zubulake v. UBS Warburg LLC*, courts introduced multifactor tests to determine when cost shifting is appropriate. In *Rowe,* the court concluded that the e-mail information sought by the plaintiffs was relevant and that a blanket order precluding its discovery was unjustified. However, balancing eight factors derived from case law, the court required the plaintiffs to pay for the recovery and production of the e-mail backups, except for the cost of screening for relevance and privilege. The eight Rowe factors were:

1. The specificity of the discovery requests.
2. The likelihood of discovering critical information.

3. The availability of such information from other sources.
4. The purposes for which the responding party maintains the requested data.
5. The relative benefit to the parties of obtaining the information.
6. The total cost associated with production.
7. The relative ability of each party to control costs and its incentive to do so.
8. The resources available to each party.

Form of production impacted by need for metadata. If metadata is relevant and discoverable, production in TIFF or PDF format could be considered incomplete or inadequate. In *Hagenbuch v. 3B6 Sistemi Electronic Industrial*, a defendant decided (against the protests of plaintiff) to convert all of the information on the original electronic media (that the plaintiff had designated for copying) into TIFF documents.

Discovery of backup tapes. In Veeco Instruments, Inc. securities litigation, the court permits search of backup tapes, rejecting argument that restoring and searching backup tapes would be unduly burdensome and costly.

High costs do not make it inaccessible. In *AAB Joint Venture v. United States,* the court ruled that several thousand dollars or tens of thousands of dollars do not make data inaccessible—requiring the government to produce e-mails from backup tapes.

Intentional spoliation. Due to its intentional spoliation of ESI, Oved Construction Services was sanctioned, had a default judgment entered against it, and had to pay its adversary's attorneys' fees. In *Echostar v. the EEOC,* the company's practice of routinely disposing of e-mails, regardless of content, was deemed "risky and extraordinary," and Echostar was sanctioned for failing to preserve e-mails relevant to a former employee's EEOC claim.

Failure to respond in a timely manner. A federal court in New York found that Strategic Resources was grossly negligent because it failed to timely produce 25 gigabytes of data, even though no evidence was destroyed.

Undue burden. In *Ameriwood Industries, Inc. v. Paul Liberman,* the court ruled that providing that information is not reasonably accessible is satisfied by showing the efforts involved in copying a hard drive, recovering deleted information, and translating recovered data in searchable and reviewable format. But a defendant is not relieved of duty to produce records merely because they chose to preserve the evidence in a format that makes the ultimate production expensive.

Clawbacks in a timely manner. In *Kuest Corp. v. Airtrol, Inc.,* the courts denied clawback because the defendants were not timely in making their claims—less than three months in this case.

Sanctioned for data preservation failures. In *Zubulake v. UBS Warburg,* sanctions were imposed on the defendant for failing to preserve e-mail. In imposing sanctions the court ruled defendant's counsel failed to communicate the litigation hold order to all key players. They also failed to ascertain each of the key players' document management habits. By the same token, UBS

employees, for unknown reasons, ignored many of the instructions that counsel gave. This case represents a failure of communication, and that failure falls on counsel and client alike.

Paying for added discovery costs from poor due diligence. In Bristol-Myers Squibb securities litigation, class action plaintiffs agreed to pay for paper copies of documents that, unknown to them, were available in a less expensive electronic format. Litigants should be careful not to place a cart blanche order for something without knowing what is available and what potential cost may inhere. Conversely, the responding party has some responsibility to explain what is available and to present reasonable alternatives to the requesting party.[9]

Limits on the scope of discovery. In *Sallis v. University of Minnesota,* the plaintiff had sought university-wide discovery of the latter's central database. In affirming the denial of the request, the court of appeals ruled that Sallis's discovery requests had no limitation—he sought information on every allegation of discrimination against the university—by all complainants in all departments. However, Sallis had spent the past 10 years working in just one department, and his allegations of discrimination focus on the behavior of the supervisors there. The court found Sallis's request to be overly broad and unduly burdensome and limited discovery to one relevant department.

Sampling discovery to determine reasonable limits. In *McPeek v. Ashcroft* and *Hagemeyer v. Gateway Data Services,* the court supported the use of sampling to tailor the scope of further discovery. The requesting party may need discovery to test the assertion that the information is not reasonably accessible. Such discovery may involve taking depositions of those knowledgeable about the responding party's information systems, some form of inspection of the data sources, and requiring the responding party to conduct a sampling of information contained on the sources identified as not reasonably accessible. Sampling of the less accessible source can help refine the search parameters and determine the benefits and burdens associated with a fuller search.[10]

Form of production impacted by need for metadata. In *Hagenbuch v. Sistemi Electronic Industrial,* the court ruled that if metadata is relevant and discoverable, production in TIFF or PDF format could be considered incomplete or inadequate. The defendant had decided, against the protests of plaintiff, to convert all of the information on the original electronic media into TIFF documents. In granting the plaintiff's motion to compel production of the information in native format, the court reasoned that the TIFF documents do not contain all of the relevant, nonprivileged information contained in the designated electronic media, such as the creation and modification dates of a document, e-mail attachments and recipients, and metadata.

Failure to follow data retention policies. In *EEOC v. Target Corporation,* the court cited 29 C.F.R. 1602, which requires employment applications for nonhires to be retained for one year. Target included this requirement in its records retention policies. The responsible Target manager was trained on the policy but failed to follow the policy, trashing them. Even though the

court ruled the case had no merit, it was reversed because the deleted e-mails could have made the plaintiff's case.

Discovery of backup tapes. In Veeco Instruments securities litigation, the court permitted the search of backup tapes, rejecting argument that restoring and searching backup tapes would be unduly burdensome and costly.

Reasonably accessible data. In *Disability Rights Council (DRC) of Greater Washington v. Washington Metropolitan Transit Authority (MTA)*, the plaintiff, DRC, alleged that the Transit Authority failed to stop its e-mail system from deleting all e-mails older than 60 days even two years after the lawsuit was filed. In its defense, the Transit Authority cited new Rule 37(f), which established a safe harbor provision for any electronic data lost as a result of the "routine, good faith" operation of an information technology (IT) system. The court ruled the failure is indefensible, finding that the Transit Authority did not act in good faith when it continued to destroy the e-mails after the lawsuit was filed and that good cause existed to require the search and production of data from the Transit Authority backup tapes.

High cost of discovery does not make it inaccessible. In *AAB Joint Venture v. United States,* the court ruled that several thousand dollars or tens of thousands of dollars do not make data inaccessible. In the case, the court required the government to produce e-mails from backup tapes, reasoning that the $85,000 to $150,000 processing cost was a drop in the bucket in light of the $30 million at issue. Specifically, the court reasoned that the government could not be relieved of its duty to produce those documents merely because defendant has chosen a means to preserve evidence which makes ultimate production of relevant documents expensive.

U.S. RULINGS IMPACTING BUSINESSES OUTSIDE THE UNITED STATES

In a variety of cases, U.S. courts have ruled that any foreign company enjoying the benefits of doing business in the United States shall fall under U.S. laws. European Union (EU) courts have backed up these decisions even to the point of waiving EU privacy protections. Here are some case law examples where foreign parent companies were not able to claim exemptions under EU privacy directives:

- *Afros, SpA v. Krauss-Maffei Corporation.* A U.S. court ruled that the U.S. subsidiary had requisite control over documents held by parent when the subsidiary was wholly owned, key litigation decisions were made by parent, and there was a substantial intermingling of management employees and directors.
- *Alcon International Limited v. S. A. Day Manufacturing Company.* A U.S. court granted the defendant's request to compel production of documents in possession of the plaintiff's German affiliate to depose employees of the affiliate. The court held that the plaintiff had control over the documents of the foreign affiliate because the two entities were corporate members of a unified worldwide business "under common control."

- *Columbia Pictures v. Justin Bunnell.* Columbia, Disney, Universal, Warner, Paramount, and other major studios sued Netherlands-based web site over for copyright infringement. Court rejected defendants' claims of protection under Dutch/EU privacy laws. There is also a random access memory (RAM) hot potato issue in that for the first time a judge ruled that a computer's RAM is discoverable.
- *Reino De España v. Am. Bureau of Shipping.* A U.S. court granted spoliation inference where the Spanish plaintiff produced merely 62 e-mails, despite extensive use of e-mail by the plaintiff.
- *Strauss v. Credit Lyonnais, S.A.* U.S. court ruled that Crédit Lyonnais, a subsidiary of France's largest financial institution, Crédit Agricole, must defend itself in U.S. court against claims by 25 families of American victims of terrorist attacks in the Middle East rejecting claims of personal privacy protection under French and EU laws and directives.

BEST PRACTICES AND NEXT-GENERATION TECHNIQUES

There are some basic best practices that organizations of all sizes and levels of complexity can follow to reduce their financial risk exposure in litigation and legal discovery. These same suggestions will help any organization in improving its ability to comply with regulations while reducing operational costs. We also offer some more advanced and next generation techniques.

Best Practices

- *Implement an enterprise-wide records and document management system.* Commonly referred to as an enterprise content management (ECM) system, they provide the proof of compliance and are the key to preventing litigation and winning cases that cannot be avoided.
- *Attack the number of siloed and disparate data repositories.* In many organizations this is a major task due to ongoing mergers and acquisitions and complicated by the lack of standardized naming and classification methodologies.
- *Federate content management.* Given the hybrid nature of document and records management in most organizations, it is essential to implement a means to federate content. This requires that links are established among the records across all repositories so searching and retrieval are truly enterprise-wide.
- *Enforce document retention and destruction policies.* This is essential given the high costs of discovery and that there many examples in which a majority of documents retrieved in litigation were retained beyond their retention requirements. A large portion of legal discovery costs are avoidable by destroying paper and electronic documents that are retained just in case. John Bace, vice president at Gartner Research, has noted:

> *Once required storage time for a record has expired, get rid of it. . . . The information quite often develops an inverse negative value. Some people say we'll keep everything forever. That is one of the worst ideas around, especially given the penalties and issues around the new discovery rules."*[11]

- *Implement workflows that control all electronic documents.* The technology has been available for many years and available for even the smallest organizations. End-to-end work or process flows automated processes and approvals while providing a transparent audit trail.
- *Inventory ESI systems and data sources.* This is a key requirement in successfully preparing for litigation and legal discovery. It includes:
  - Content (what types of records and data).
  - Custodian (the owner or system administrator).
  - Location.
  - Preferred form or format of production.
  - Initial assessment of cost and burden of production.
  - Initial assessment of privileged data.
- *Prepare for litigation holds.* Courts are imposing severe sanctions for failing to comply with litigation holds. Preparation should include:
  - The creation of a documented litigation hold policy and procedure.
  - Record custodians and system administrators understand their roles.
  - Litigation (actual or reasonably anticipated) triggers notices to custodians to suspend disposal. Confirmation, both initially and periodically, is required to validate the process.
  - A plan for collection of ESI from global sources is well understood and in place.

Next-Generation Techniques

- *Digitize and classify all documents upon creation.* In a born-digital environment, the process is automated and paper originals are viewed as a liability.
- *Destroy all paper documents unless required by regulations.* There are few valid examples in which original paper documents are still required. Scanned and digital signatures are widely accepted by most all regulatory agencies.
- *Implement complex federated content management.* The goal is to cross-reference and make all related records and documents readily and cost-effectively available for searches and analysis. All documents related to a given customer or supplier would be federated, or linked. In the example of a bank, this would translate to all loan, savings, checking, money market, and retirement accounts. Complex federated content management should also include the ability to perform deep data mining, search and retrieval of manageable amounts of electronic documents and metadata. Manageable means reducing the number of false positives, which are the bane of all litigants, auditors, and regulators.
- *Do battle with your own legal department.* Legal departments will tend to want to keep everything forever as a just-in-case defensive strategy. Too much data is an expensive liability, but legal departments have little incentive to reduce the burden on IT and business owners in a legal discovery process. This will require support at the highest executive levels and outside legal advice to reduce mountains of paper documents that are so costly to maintain and so difficult to recover. The key is to challenge the assumption that lots of paper is good. To the contrary—it tends to be a liability, and rarely is it required to keep paper originals. What should remain for documentation should be in an

electronic/digital format in which all related documents, records, and metadata are cross-referenced and easily recoverable and searchable.

CONCLUSION

Litigation and legal discovery are now a major headache for most American firms. The pain is spreading to America's trading partners as well. As it does, it will also change their nature from more conciliatory and dispute resolving to one where litigation becomes the first choice. The United States was not nearly so litigious in the past, and there is nothing on the horizon to suggest it become substantially less litigious in the coming years, and, to the contrary, the problems will spread beyond U.S. borders.

Some reforms are being discussed to reduce the size of punitive damages. There is almost virtual unanimity in most all countries that a damaged party should be made whole—covering their direct costs. But U.S. juries have used punitive damages to cover pain and suffering and also to punish large and unpopular corporations. U.S. litigants have also used the courts to cover the gaps in regulatory enforcement and corporate governance.

Maybe the best hope to reduce litigation is to improve corporate board governance which we detail in Chapter 24. We call for increasing minority representation on boards, separating the roles of chief executive officer (CEO), and chairman of the board (CoB), and creating a risk committee run by risk experts reporting directly to the board. These measures will tend to make companies more prudent and thoughtful in making business decisions that foster disputes with their competitors, customers, suppliers, and employees. There is a need for tort law reform to lower the abuses in punitive damage awards and the filing of the vast numbers of law suits.

Since relief is, at best, years away, organizations must be prepared to face a very costly response process and losing lawsuits that can substantially damage reputations and financial viability.

NOTES

1. John Bace, "Cost of E-Discovery Threatens to Skew Justice System," Gartner Research, April 20, 2007.
2. Business Wire, "Litigation as the Great Equalizer: New Fulbright & Jaworski Survey Finds Nearly 90% of U.S. Corporations Engaged in Lawsuits; Average $1 Billion Company in U.S. Faces 147 Cases at a Time," *Business Wire* (October 10, 2005).
3. The Sedona Conference®, "The Sedona Principles," 2nd ed., Best Practices Recommendations and Principles for Addressing Electronic Document Production, July 2007.
4. The Sedona Conference®, "Glossary For E-Discovery and Digital Information Management," May 2005.
5. The following link contains the new rules: www.supremecourtus.gov/orders/courtorders/frcv06p.pdf.
6. Timothy Carroll and Bruce Radke, "The Amendments to the Federal Rules of Civil Procedure Concerning eDiscovery Impact on Global Business Enterprises, Busmanagment.com.

7. Barbara J. Rothstein, Ronald J. Hedges, and Elizabeth C. Wiggins, "Managing Discovery of Electronic Information: A Pocket Guide for Judges," Federal Judicial Center, 2007.
8. See: The eDiscovery and Analysis Group, "Electronic Discovery Law Blog," K&L Gates, http://www.ediscoverylaw.com/articles/case-summaries/.
9. A. Blakley, ed., "Electronic Information 62-63." (Federal Bar Ass'n: 2002).
10. Barbara J. Rothstein, Ronald J. Hedges, and Elizabeth C. Wiggins, *Managing Discovery of Electronic Information: A Pocket Guide for Judges* (Federal Judicial Center, 2007).
11. John Bace, VP of Research at Gartner (Compliance Week, October 16, 2007).

CHAPTER 17

The Circle of Trust

Brett Trusko

INTRODUCTION

In the 2000 movie *Meet the Parents,* Robert De Niro's character is an ex-CIA agent who keeps reminding Ben Stiller's character that since he is joining the family he is in the "circle of trust" unless he "blows it."

This chapter discusses something that is completely hypothetical and to our knowledge is not done by any trading or banking partners as of today. Instead, this is a discussion of what might be accomplished in an arrangement where two partners cooperate to create excellence in their business relationship.

The circle of trust will be discussed in reference to a Six Sigma program and a new paradigm whereby organizations agree that inspection would be eliminated if excellence could be achieved, reported, and verified. So, as an example, let's suppose that a health insurance company receives and processes insurance claims for a major hospital. The current process is that all claims submitted are verified either manually and/or electronically against certain benchmarks and/or other verification processes. The working paradigm is that all claims have errors or intentional misrepresentations and therefore must be audited before paid.

In a mortgage market, let's assume that we are a financial services company purchasing packages of mortgages from a bank. As we have recently discovered, there has been little oversight in purchasing these bundles of loans due in part to the cost, and also the statistical risk that most of these loans are good and that the cost of verifying the loans is less than the risk, since in many cases these were simply repackaged and resold anyway. A viable strategy until the financial musical chairs music stopped and organizations found they were without a chair.

How could this have hypothetically been addressed by our circle of trust? Consider the following assumptions concerning Six Sigma:

- Six Sigma is generally considered impossible to reach (with exceptions), but in certain circumstances, there is nothing wrong with three sigma. In fact, if pay is linked to performance and incremental improvements in pay are linked to incremental improvements in sigma, then the prime motivator is the sigma level itself and not necessarily the number of errors per million.
- Improvements in sigma level can be assumed to have economic value. If the value of a portfolio of loans can be considered higher as number of errors in the

loan are reduced (statement of earnings attached, credit score verified, property inspection reports, etc.), due to a lower risk contained in the portfolio then can we also assume that the value of that portfolio is higher with the reduced risk?
- Traditional accounting firms or specialized auditing entities that specifically audit sigma levels for a finite business process (loans, insurance bills, etc.) for conformance to defined sets of rules that can then be translated to a sigma level.
- Contracts with sliding scales for high performance can be written and will be honored.

To date, we are not aware of anyone utilizing Six Sigma as a contracting tool, but as we will discuss in the remainder of this chapter, the circle of trust is a viable option in financial services and insurance transactions that can allow for excellent performers to be paid more and the quality of portfolios of anything from insurance transactions to home loans to stock transactions and others is within the imagination of the business partners.

IS THREE SIGMA GOOD ENOUGH?

Three sigma is defined as approximately 66,800 defects per million opportunities. This is traditionally thought of within the Six Sigma community as a relatively poor performance level, but when taken in the context of the subprime mortgage problems in the United States, where some estimates are that up to 58 percent of all loans in 2006 may contain documentation errors, according to First American Loan Performance.[1]

Note that although subprime loans (sometimes called liar loans) are not necessarily an issue of "error" since in fact much of the problem with the subprime market were in fact loosening of the rules versus flat out lies. Many of the bad loans were due to a rush to loan. Fraudulent W-2s (proof of income), lack of proper counseling by mortgage brokers who were rewarded for convincing borrowers to take more expensive loans, and setting of unrealistic expectations by lenders contributed to the subprime loan crisis.

In our second case of insurance claims in a health care environment, the question is one of "gaming" of an insurance claim to achieve the highest reimbursement level as opposed to actually trying to state the truth. In the insurance game, particularly in health care insurance, there is absolutely no trust between the two sides. The health care provider (although most would deny it) actively tries to increase their reimbursement while (also with deniable plausibility) the insurance company tries to slow reimbursement, via mechanisms of review and audit, (averaging in the mid-50-day range in 2007 according to the Healthcare Financial Management Association) in an attempt to pay the correct amount. Additionally, as has been communicated extensively, up to 30 percent of the cost of health care is in administrative overhead, with a large amount of this cost coming from the claims administration process.

Would some assurance that claims were filed accurately allow for a circle of trust with the insurance company and lead to faster and higher reimbursement? With some assurance of a strict set of predefined quality criteria, the foundation of a circle of trust relationship could be the foundation for elimination or reduction of administrative overhead in this business relationship.

One of the biggest strengths and in some respects weaknesses in Six Sigma is the dogged adherence to the notion of the "Voice of the Customer" in the definition

of quality. In the case of the circle of trust, it may not be readily apparent who the customer is, and in fact in many cases the customer is bilateral and both sides of the transaction can benefit from improvements in quality of the transaction. After all, in the subprime mortgage transaction, the customer can be difficult to identify. In the case of a circle of trust in a health insurance transaction, the relationship becomes symbiotic and therefore both parties are the supplier and the customer. Significant negotiations defining the measurements of quality and the rewards are critical for the circle of trust to work. Also, perhaps by utilizing the "measurement" of Six Sigma versus the "program" of Six Sigma, the program could be an evolutionary enhancement to improve sigma levels and in turn improve the relationship and economic value.

ECONOMIC VALUE OF A SIGMA

One of the mainstays in the Six Sigma movement is the cost of poor quality. We know that the cost of poor quality is the economic cost of what the organization spends to make up for its mistakes/errors. This cost is traditionally thought of from a "lost opportunity" perspective or what it will cost the organization if they make a mistake/error. This has never been argued to be easy. If, for example, a physician fails to take an x-ray, then what is the cost of a malpractice suit? If a mortgage broker fails to verify income, what is the chance that someone will default on their mortgage? And the most famous error of all time:

For want of a nail, the shoe was lost;
For want of the shoe, the horse was lost;
For want of the horse, the rider was lost;
For want of the rider, the battle was lost;
For want of the battle, the kingdom was lost.

Of course, anyone seriously considering this statement realizes that although it may have been the nail, nothing is as black and white as a single nail. For example, why didn't we have a nail? Why wasn't the nail installed properly, and so on?

What we can do and would work well in the circle of trust is to establish the probability of the missing nail leading to the loss of a battle. While the example of the want of the nail is a bit silly, it is an easier illustration of how we can apply the circle of trust than with something much more difficult to follow.

First, what was the probability that a nail would be lost? If we look at the probability of a shoe being thrown, then we have a good place to start. While I couldn't find any studies in the probability of a horse throwing a shoe, let's assume that the probability is somewhere around 1 in 500 each time a rider gets on a horse, and that the reason for that shoe's being lost is typically that the shoe wears out and not because of the loss of a nail. Now the odds are probably somewhere around 1 in 10,000 for a single nail, one in 4,000 for 2 nails, 1 in 500 for three nails, 1 in 25 for 4 nails, and of course five missing nails assures you that the shoe will not stay on.

Now, let's go to the loss of the kingdom. The king has to be asked first, why are we fighting the battle anyway? Is it because the king levies too many taxes on his subjects? Was the battle fought in a strategically poor position? Were there too few bullets available to his soldiers? And why were there too few bullets? Was the bullet

shortage due to poor supply chain management that also led to too few horse shoe nails? Did we decide to create more bullets at the expense of nails because we only may have thrown a shoe and lost a battle, but we would have almost definitively lost the battle without the bullet? And was it one bullet that this cost us or was it thousands? Did we also decide to make a decision about shoes whereby all the horses only receive four nails, since the odds are very small for leaving out one nail.

This ridiculous discussion is one that would almost certainly not be made in today's world of electronic communications, but in the real world they are the kinds of economic decisions that are made each and every day. Yes, every horse should have all five nails, and in reality no horse should leave the stable without all five nails, but if the cost of getting five nails on every horse is the failure to produce enough bullets, then perhaps putting more resources on a bullet manufacturing line is the way to go.

Now, applying real-world examples to our problem, let's discuss the economic value of a sigma. Let's assume that the average widget application is filled with information. The widget is a highly customized service, and no two widgets are exactly alike (this could be a home mortgage, a car loan, or an insurance claim). Now let's assume that there are two risks to the organization paying for the widget: the risk of incomplete information causing regulatory issues and the risk that they pay too much for a widget if they don't fully understand what widget was actually delivered.

Now, since the risks are high that a widget is either illegal or that they may pay too much for a widget they didn't get, the organization in question has made the decision that they will review (manually or by computer) all widget transactions, at a very high cost to the organization. Now, let's assume that some very bright young employee declares that they practice of 100 percent inspection of a widget is ridiculous. The corporate accountants declare that no inspection is just as ridiculous. Are they at an impasse? And what percent of widget bills/applications is the right number? The solution might be the circle of trust.

Imagine that the company that supplies the widget and the company that pays for the widget get together and agree to certain performance standards. Yes, this happens in today's business world, but what if there were independent third parties to confirm that the standards were being met. In many cases today, there are go/no go measurements of performance (e.g., in a bond covenant), whereby something happened or didn't, a certain profit margin was or was not met, and so on.

In the circle of trust, there would be a scale such as Exhibit 17.1.

So a sigma can have economic value if removed from the Six Sigma process, but if the process is employed it is already something a good Six Sigma company should be capturing.

THE SIX SIGMA AUDIT

The circle of trust is fine as stated in the movie, but in the real world, as Reagan liked to say, "trust but verify." For companies that do business together to unconditionally trust each other may happen in Utopia, but in the real world of business, trust breaks down quickly when an error, omission, or deception takes place. In the real world, we have audit firms and attorneys.

EXHIBIT 17.1 Circle of Trust Scales
We pay "X Dollars" for a widget . . .

2 Sigma	We are required to do extensive audit of widget bills so we pay *X* minus 30%.
3 Sigma	We can now do statistical audits of bills, since we know that most are correct. We will pay *X* minus 15%.
4 Sigma	We know that there are very few errors in the bill, we now feel comfortable paying *X*.
5 Sigma and above	Since we now do not need to worry about regulatory issues, because the widget bills are almost always accurate. Not only do we not need to worry about overpaying, we are also mitigated against regulatory loss and therefore we can pay a bonus because we save money on mitigation. We will pay *X* plus 10%.

So how would one work with the circle of trust? Quite simply, utilize existing infrastructure of audits and contracts. So the sigma audit might look something like this:

- Attorneys for both sides would partner with process experts to determine what a perfect process might look like in the stated relationship. The process should be directly related to the business relationship being discussed. For example, it would be appropriate to demand performance in correctly and accurately collecting loan data and documents, but it would probably be inappropriate to audit the completion of training requirements for human resources personnel, since the latter requirement is only ancillary related to the requirement that a document be completed accurately and that all documents related to the loan are collected and contained in the files.
- The process engineers and/or quality experts will be required to create an analysis of, instead of the cost of poor quality, the economic value of outstanding quality. This analysis would be the basis for negotiations of sigma levels in the creation of the base rate for the contract.
- After completing the negotiations of the process capability and the economic value of the process, the sigma levels would be negotiated and on a regular basis (agreed upon in the contract), a statistically basis audit of the process would be done to determine payment levels for the next period. In addition, something akin to an audit report would be prepared by the auditing entity reporting the findings of the audit. This would aid the business partner because they would then know which areas of the transaction would need to continue to be reviewed. This would allow the partnership to continue to review areas that are not well done, while dismissing areas that are done well.
- Payments would be adjusted periodically based on the requirements of the contract.

Benefits to this approach may not appear readily apparent without additional discussion. Primarily, if there were enough organizations working with a circle of trust the overhead related to review of transactions would be virtually eliminated.

Perhaps not with the first circle of trust, but as additional circle of trust partners are added a single point of contact verifies the quality of the transaction. Therefore, the greater number of organizations that are participating in the circle of trust the lower the cost to the entire system and to society as a whole.

A second benefit is that organizations would become more focused on improving the quality of their business. In macroeconomic theory, this benefits the entire industry. Additionally, as we discovered in the subprime mortgage crisis, we have the basis for assuring ourselves that when we purchase a portfolio of loans that documents are included, incomes verified, credit scores are accurate, and disclosure documents have been discussed and filed truthfully. Could this type of system have reduced the impact of the subprime crisis? We will never know, but it could be interesting to contemplate.

A third and final benefit is that the cost of overhead from review of all or a large part of transactions could be eliminated or at a minimum reduced significantly. In the case of health insurance claims, it is widely publicized that up to 30 cents on every health care dollar is spent on administrative overhead. Much of the overhead is related to verification of claims by humans. If claims submitters and health insurance companies entered into a circle of trust relationship, the health insurance company could save up to $399.4 billion per year,[2] which, if split between business partners, would allow the health care system and insurance companies to realize greater profits while reducing the cost of health care to employers and patients.

CONCLUSION

The need to improve quality in financial transactions coupled with the need to reduce costs in an ever more competitive environment and increasing compliance requirements by regulatory agencies demand more cooperation between business partners. While these organizations are generally prohibited from directly coordinating their efforts, there is no rule against trusting your business partners. In fact, many companies explicitly trust their business partners, but these relationships are due to years of working together. Other business partner relationships reduce costs by employing strong arm tactics such as Wal-Mart and other large retailers. For smaller and less aggressive partnerships a few dollars invested in the circle of trust will reap great benefits almost immediately.

NOTES

1. NPR Radio News, August 7, 2007, www.npr.org/templates/story/story.php?storyId=12561184.
2. "USA Wastes More on Health Care Bureaucracy Than It Would Cost to Provide Health Care to All of the Uninsured," *Medical News Today* (May 28, 2004).

CHAPTER 18

Reducing Liability Risk through Best Environmental Practices

Nasrin R. Khalili, Ph.D.

INTRODUCTION

Most corporations have recognized the inevitable need to manage environmental risks associated with the industrial economy. They have also recognized the necessity for mitigating this risk in order to succeed in the competitive market. Mitigation strategies, however, deliberate on both the economic and environmental concerns within the dominion of corporate strategy. Historically, economic and environmental goals have been perceived as divergent forces with the perception that economic criteria must be satisfied before environmental goals are pursued. The concept of sustainable development, however, argues that the economics and the environmental goals are neither mutually exclusive nor necessarily conflicting. We thereby present here a comprehensive analysis of environmental risk mitigation strategies through application of environmental technology at selected leading organizations.

The results of the analysis indicated that in addition to traditional criteria, such as cost, quality, and performance, environment is becoming a critical operating component in both product realization processes and operations. It has also been observed that including environmental decision making throughout the entire value-adding processes could result in synchronous economic and environmental growth. Successful integration of environmental values into the operation context, however, should involve both recognition and renovation of a complex mix of interacting factors. For example, continuous environmental improvement, which is also economically beneficial, can be achieved by enhancing efficiency and productivity of industrial systems, and maintaining material use, recycling, and reuse in the industrial context just as techniques for proficient use of materials and energy are explored. Efficient use of cross-functional teams, better understanding of the product line and operations, and innovative supply chain management practices are all linked to proactive environmental policies, ultimately resulting in stronger environmental and economic performance.

Historically, economic and environmental goals have been perceived as conflicting forces. Studies focused their direction toward investigating the trueness of the common believes that economic criteria must be satisfied before environmental goals are pursued. The link between environmental and economic performance has been

widely studied in the literature. The common belief is that improved environmental performance is responsible for extra costs to the firm and so it reduces profitability. The counter opinion, however, is that improved environmental performance would induce cost savings and increase sales and improved economic performance. Schaltegger and Terje Synnestvedt showed that the main factor impacting economic outcome of the firms is not the level of environmental performance, but the kind of environmental management with which a certain level of success is achieved.[1] Accordingly, they suggested that research and business practice should focus less on general correlations and more on causal relationships of eco-efficiency.

The conception of sustainable development, however, argues that the goals of environmental and economic growth are neither mutually exclusive nor necessarily conflicting. A comprehensive analysis of the environmental risks mitigation strategies presented in this paper depict the concord of economics and environmental growth by means of including environmental decision making into the entire corporate value-adding process. As presented in the following sections, successful integration of environmental values into operation context calls for both recognition and renovation of the existing complex mix of interacting factors. Accordingly, one must examine those factors and the apprehension to including environmental concerns within industrial economy. Elements of industrial environmental management and the impact of environmental regulations on firms' profitability and competitive advantage are discussed by presenting Kuznets environmental curve concepts and the Porter hypothesis that environmental regulations result in competitive advantage via innovations.

Studies suggest that depending on the dynamic or static nature of the managerial perspectives, firms presume that the impact of environmental practices and regulations on industrial performance (profitability and growth) will be either conflicting or complementary. It is believed that continuous environmental improvement that is also economically beneficial can be achieved by enhancing efficiency and productivity of industrial systems.

After decades of end-of-pipe treatment and controls on industrial releases to the environment, attention has shifted to including elimination of potential pollution at its source, design with environmental factors in mind, and sustainable manufacturing. These efforts have incorporated engineering attempts to redesign products and processes, incentives to encourage pollution reduction, and pollution prevention while focusing on sustainable development. Many studies in the early 1990s showed that appropriate environmental policy and government regulation are the most important catalysts in leading firms to consider environmental issues today. Forces such as customer pressure, shareholder pressure, and minimizing financial and social risks may also play a significant role in the development of an environmental plan at the firm level. Since various empirical studies suggest that most firms already spend between 1 and 2 percent of their revenues as a response to environmental concerns, it is becoming increasingly essential for firms to develop a firm corporate environmental policy.

The mission of sustainable development is to equilibrate economy with resources and natural ecosystems. Sustainable development is a concept that requires restructuring of social, economic, technological, and industrial policies and practices. Sustainability goals can be achieved through environmentally desirable changes (EDC) in industrial production, that is, eliminating waste, changing production processes,

redesigning products, fostering profitable innovation, and promoting energy conservation. Decades of review of industrial performance suggest that firms can gain competitive advantage from redesigning production processes to be less polluting, substituting less polluting inputs, recycling by-products of processes, and instituting less polluting processes. Such approaches reduce the cost of production by increasing the efficiency of production processes and reducing input and waste disposal costs.

THE ECONOMY AND THE ENVIRONMENT

The relationship between the economy and the environment has been focus of many studies. Materialist and postmaterialist approaches suggest that economic needs must be satisfied before environmental goals are pursued. The sustainable development perspective, however, stresses that the economic and environmental goals are neither mutually exclusive nor necessarily conflicting.[2]

Theories of "motivation" illustrate reasons for pursuing green production and sustainability. Maslow's theory of motivation (1970), suggests that economic fulfillment is a necessity while environmental concerns are higher needs related to association and the quality of life. Accordingly, societies pursue basic needs such as economic satisfaction before considering higher goals such as environmental protection. Inglehart, expanding on the work of Maslow, hypothesized that societies pursue goals in hierarchical fashion. Once more basic needs are satisfied, generations pursue the ordained higher needs. Inglehart found that the more basic, materialist values are those of physical sustenance and safety, while the higher, postmaterialist values are those of quality-of-life concerns, including environmentalism.[3]

Other perspectives, however, do not pose economic and environmental goals in conflicting, win-lose opposition. The popular conception of sustainable development, for example, presumes that economic development can occur in an environmentally benign way. The World Commission on Environment and Development has identified the gaps and inequities between industrial and developing countries as the core of both environmental and development problems, and suggests exploring solutions that promote economic growth that is equitable, and environmentally sustainable.[4]

The Kuznets curve theory (Simon Kuznets, 1955, 1963; and Grossman and Krueger, 1993, 1995) suggest an inverted U-shaped relation between income inequality, economic growth, the size of an industry/economy and environmental pollution. According to Kuznets's theory, as income increases, pollution also increases to a point (win-lose situation), after which it decreases with increase in income (win-win situation).[5] The Environmental Kuznets Curve (EKC) suggests an approximated link between environmental change and income growth. The most popular indicators of the EKC are the inverted U-shape curve found between local air pollutants and per capita income.[6]

Despite many findings, including studies relying on the application of sophisticated econometric techniques to explain this theory, there is still no clear-cut evidence to support the existence of the EKC.[7] In another study, Grossman (1995) identified three main channels whereby income growth affects the quality of the environment; larger-scale economic activity leads, per se, to increased environmental degradation. This occurs because increasing output requires more inputs and thus more natural

resources being used in the production process. Higher output also implies increased waste and emissions (by-product of the economic activity). Yet income growth can also have a positive impact on the environment through a composition effect so as income grows, the structure of the economy tends to change, gradually increasing the share of cleaner activities in the gross domestic product.[8,9,10]

The growing connection between the environment and economic growth has created many challenges for business. In response, a set of recent dialogues, convened by the Aspen Institute, focused on the business opportunities inherent in environmental leadership. Their analysis suggested that businesses that integrate their environmental planning with their strategic business planning can improve their corporate performance and gain a competitive edge. They also suggested that investors and analysts who understand these connections will be better positioned to identify companies with superior stock appreciation in the newly emerging sustainability driven marketplace of the twenty-first century.[11]

ENVIRONMENTAL RISKS: RISKS AND THE SECURITIES AND EXCHANGE COMMISSION (SEC)

Corporations often have real environmental liabilities that translate into both harm to the environment and harm to shareholder value. By requiring corporations to disclose potential liability, shareholders will get a better idea of how such activity could influence the likely direction of their holdings.

In 1998, the Environmental Protection Agency (EPA) Office of Enforcement and Compliance Assurance completed a study that reported significant under disclosure of corporate environmental liabilities. Among other findings, it revealed that 74 percent of companies failed to comply with the U.S. SEC regulations governing the disclosure of environmentally related legal proceedings that could result in sanctions exceeding $100,000. The SEC had "no comment" on the EPA findings. In 1990, five insurance companies stated that they were involved in potentially costly environmental claims that could have negative financial impacts on their bottom lines, but only two of these companies disclosed the dollar amounts of these claims. In 1991, eight companies admitted such environmental claims, but only three disclosed dollar amounts in their annual reports.

In 1993, upon reviewing 16 insurance company annual reports (The Environmental Liability Report : Property and Casualty Insurer Disclosure of Environmental Liabilities), the General Accounting Office (GAO) released a report indicating that very low levels of insurance company disclose Superfund toxic cleanup liabilities.

Following disclosure of these reports and the reviews, on August 21, 2002, The Rose Foundation filed a petition with the U.S. SEC proposing a new rule to govern corporate disclosure of environmental liabilities. Upon submission of the petition, more than 20 environmental and community foundations representing over $2 billion dollars in combined assets sent a letter to SEC Chair Harvey Pitt on this topic. In conjunction with these efforts, the Rose Foundation also released a report documenting how corporate environmental liabilities can impair shareowner value. In the latest report, Tim Little, cofounder of the Rose Foundation, claimed that despite the recent accounting reform prompted by the Enron and WorldCom scandals, American

investors are still at risk to lose their hard-earned savings because corporations cook their books by keeping environmental costs off the balance sheet. Both the report, entitled "The Environmental Fiduciary: The Case for Incorporating Environmental Factors into Investment Management Policies," and the petition hinge on two key U.S. government findings. In addition to documenting corporate underdisclosure of environmental liabilities, The Environmental Fiduciary also has presented substantial evidence of a positive correlation between financial performance and environmental performance.

The insurance companies claimed in the GAO report that the lack of guidelines and rules for estimating potential costs of environmental liabilities prevented them from disclosing this information. The insurance industry responded to the report by contracting the American Society for Testing and Materials (ASTM) to develop a set of guidelines for environmental disclosure.

After a seven-year, full-consensus process based on industry input, the ASTM proposed a protocol for disclosing environmental liabilities. The Rose Foundation proposes employing the ASTM's guidelines as the template for a new SEC rule governing disclosure of environmental liabilities.[12]

Due to a wide range of uncertainties, often it is difficult to measure the environmental risks. The uncertainties in its extent, timing, and the definition of terms, to mention a few, have contributed to the difficulty, calling for development of environmental risk quantification tools as an aid in a manager's decision-making process. Environmental risk rating (ERR) should be developed in such a manner that it can give a reliable and predictive link between environmental and financial performance. The rating could be used by the financial sector to set terms and pricing of products, by discriminating between companies' environmental performance.[13]

In a study conducted by Kennedy in 2001, a new series of templates were proposed to assess level of environmental risks. In this study, each template consisted of five boxes each with a description of some of the attributes of an environmental management system. The boxes were numbered one to five with each box representing an increasing level of sophistication of the system. The system was refined using a second set of templates to measure environmental risk. A matrix in which each of the environmental disciplines was plotted according to its risk and the management systems in place to control it was then developed. The procedure resulted in defining risks that were not managed properly through the system in place, therefore, assisting companies to develop appropriate environmental management systems.[14]

The chemical and related process industries are particularly exposed to high environmentally related risks and therefore costs arising from normal operation and accidents. A methodology (process environmental risk assessment [PERA]) was developed and presented for the assessment of all such risks during the design of new processes by Sharratt and Choong using a life-cycle approach while centering on risk assessment that seeks potential problems along the whole supply chain. The study mainly focused on defining activities, resource use, and/or waste source generations along the supply chain. The relevant stakeholders had to be identified prior to assessing the risk to the project or process supply chain. This study resulted in development of a methodology for facilitating management and communication of risk while applied to the manufacturing of the new product/process or the existing processes.[15]

IMPACT OF INDUSTRIAL ENVIRONMENTAL MANAGEMENT ON FIRMS COMPETITIVE ADVANTAGE

Many studies have documented environmental risk mitigation, and performance improvement through effective use of environmental management strategies. Hart suggests that capabilities that evolve as a result of a firm's response to competitive environments would influence competitive strategies and organizational outcomes. He argues that innovative environmental strategies can lead to the development of firm-specific capabilities, potential sources of competitive advantage. Their study also suggests that as firms become more constrained and dependent on an ecosystem, in order to become competitive in the market, they must develop capabilities that involve interconnected strategies for pollution prevention, product stewardship, and sustainable development.[16]

Early 1990s literature on the effect of environmental regulations on firm's capabilities has specified that appropriate environmental policy and government regulation have been the most important sources of pressure on firms to consider environmental issues in their processes and operations. Initially, many firms supposed that environmental regulations poses costs and, therefore, hinder the ability of the firms to compete in an international market. Despite this common belief, Jaffe et al. showed that environmental regulations are in fact a net positive force driving firms and the economy to be more competitive. Forces such as customer pressure, shareholder pressure, and minimizing financial and social risks may also have played significant roles in the development of environmental plans at corporate levels. Empirical studies have shown that it is becoming essential for firms to develop tangible corporate environmental policies, since most firms already spend between 1 and 2 percent of their revenues on environmental concerns (up to 1998).[17,18]

The literature provides a variety of other sources on how Industrial Environmental Management (regulations, standards, and pollution abatements) is viewed at the organizational level. Rugman and Verbeke suggested that, depending on the dynamic or static nature of the managerial perspectives, firms presume the impact of environmental practices and regulations on industrial performance (profitability and growth) to be either conflicting or complementary. Industrial performance reflects conventional parameters such as profitability and/or growth. Environmental performance, however, is defined by emission levels, degree of resource consumption, and measures of ecological impact.[19]

The challenge facing firms today is to determine whether environmental regulations set to improve environmental performance conflict with maintaining industrial performance, or complement and perhaps even improve performance. It is believed that continuous environmental improvement that is also economically beneficial can be achieved by enhancing efficiency and productivity of industrial systems.

The concept that environmental progress and competitiveness are not inconsistent but rather complementary was first proposed by Porter in 1995. In this work, Porter discusses the mechanisms utilized by firms to approach environmental practices. He argues that where firms exercise cost-minimizing choices, environmental regulation could be perceived as costs and therefore, may affect the market share of domestic companies in global markets. Competitiveness at the industry level arises from superior productivity, in terms of either lower costs than rivals or the ability to offer products with superior value that justify a premium price.

Over the last two decades, there has been a shifting focus from the static to dynamics models, industry has realized that environmental practices such as pollution prevention could positively impact the bottom line, and foster international competitiveness via innovation.[20]

Rondinelli and Berry showed that better understanding of how natural environments function is converging to provide new opportunities for environmental management that goes beyond regulatory compliance to reduce pollutants (air pollution). In cooperation with government, businesses in every industry can play crucial roles in achieving higher standards of air quality while at the same time maintaining acceptable levels of economic growth. In this study, they explored/suggested three ways in which corporations can contribute to environmentally sustainable development: (1) by adopting proactive environmental management systems that focus (in this case, on air pollution prevention); (2) by developing new technologies for air pollution control and reduction; and (3) by transferring air pollution control and prevention technologies through international trade and investment.[21]

To stay competitive, firms are realizing that environmental issues are becoming a critical operating component in most facets of industry, particularly in the manufacturing process. Along with traditional criteria such as cost, quality, and performance, environmental practices are being increasingly considered in the product realization processes and operations of companies. In assessing the main manufacturing practices, concurrent engineering and total quality management (TQM) have been targeted to facilitate the integration of environmental factors in industrial production. Using existing tools or developing new TQM programs, such as Total Environmental Quality Management, it is possible to translate environmental responsibilities across all aspects of industry, particularly in initial design activities and phases of product development. As environmental considerations are viewed as critical components of concurrent engineering practices or as important quality issues, they are more easily accepted as standard practice within companies.[22]

As indicated, firms could reduce their waste by enhancing their process performance and efficiency through TQM practices. In most industry, because of the poor flow of information, the value of TQM is underestimated, and most information about the value of defect reduction is often both delayed and masked throughout the system. The value of the waste treatment, is, however, somewhat more clear and understandable with regard to day-to-day operation since it has a static effect. As a result, most firms often miss the opportunity to implement TQM principles in the production process as they focus only on fixing quality problems at the end of the line.[23] Womack et al. argue that the difficulty of observing the true value of waste-free (lean) production and TQM practices has impacted industrial approach to sustainability. Without a clear view, diffusion of the concepts of waste minimization through TQM in industrial operations is next to impossible. Firms who understand the value of such practices, however, can improve their financial performance by enhancing process quality and reducing end-of-line quality control.[24]

The reasons for such delays were studied by Allanby. He suggests that the existing accounting systems in industry prevent firms from understanding and internalizing environmental costs and considerations correctly. Institutional barriers preventing firms from getting the information necessary to pursue optimal environmental consideration strategies are related to current accounting systems, which are not designed to encapsulate much of the engineering and accounting data required

for environmental decision making. Engineering and accounting data are usually aggregated in such a way that they lose their environmental information and managerial control. Therefore, it is logical to expect that firms must develop independent techniques for scrutinizing their product lines and operations if they wish to earn a competitive advantage by pursuing proactive environmental practices.[25,26]

Sharma and Vredenburg studied the validity of a hypothesized relationship between environmental responsiveness strategies and the emergence of competitively valuable organizational capabilities. They examined corporate environmental strategies along 11 dimensions. Utilizing two different sets of regression models they tested relationships between environmental strategy, firm capabilities, and observed benefits to firms resulting from environmental strategies. Their analysis predicted that companies that score higher on environmental responsiveness strategies will score higher on the organizational capabilities, and higher levels of competitive benefits will be associated with higher scores on the organizational capabilities measure.[27]

Claver et al. conducted a case study (COATO farming) to demonstrate relationship between environmental management and economic performance. They particularly focused on studying the relationship between environmental strategy and firm performance: the combination of environmental performance, competitive advantage, and economic performance. Their study showed that environmental management, while focused on prevention logic, had a positive net effect on the farm environmental performance. The order in which the practices were adopted resulted in development of new organizational capabilities, derived from the experience of employees who actively participated in creating new projects to reduce residues and pollution. The obtained competitive advantage was attributed to the brand image and to increased credibility in business relationships. Results also suggested a positive correlation between the pioneering proactive strategy adopted by this cooperative and the improvement of its firm performance with respect to the other firms in its sector.[28]

SHIFT IN INDUSTRIAL ECOSYSTEM TOWARD SUSTAINABILITY

After decades of end-of-pipe treatments and controls on industrial releases into the environment, attention has shifted to the elimination of potential pollution at its source, design with environmental factors in mind and sustainable manufacturing.[29,30] These efforts range from engineering attempts to redesign products and processes, incentives to encourage pollution reduction, and pollution prevention, all while focusing on sustainable development. The mission of sustainable development is to equilibrate economy with resources and natural ecosystems. Sustainable development is a concept that requires the restructuring of social, economic, technological, and industrial policies and practices. Recent environmentally friendly moves by industry and new regulatory approaches, such as emission trading and voluntarily programs, contend that firms can benefit from pursuing sustainability and environmental protection.

Sustainability goals can be achieved through environmentally desirable changes (EDC) in industrial production, that is, eliminating waste, changing production processes, redesigning products, fostering profitable innovation, and promoting energy

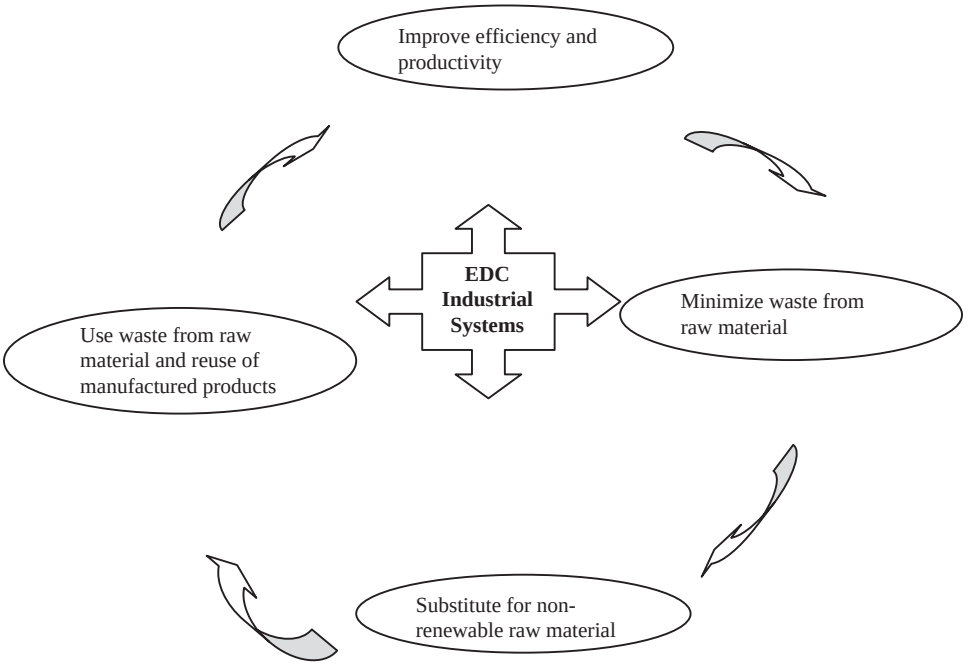


EXHIBIT 18.1 Categories Identified for Potential EDC

conservation. A survey of the current ecology of industrial systems suggests several potential EDCs in industrial production and practice. Exhibit 18.1 illustrates categories identified for potential EDC.

As shown, efficient materials and energy use requires technological and managerial innovation that would result web of waste recycling and reuse found in natural systems be imitated in the industrial context. Allanby et al. delineated three stages for the industrial ecosystem. At one stage there exists a system employing a one-way flow of materials and energy, stipulating production, use, and disposal of products to occur without reuse, or recovery of energy or materials (current industry without responsible corporate environmental programs). Additionally, there exists a system that utilizes some internal cycling of materials, requiring some virgin material input, and treatment and control of generated wastes outside the economic system. Third, there is a hypothetical system with zero discharges; most similar to the natural systems, it would involve complete or nearly complete internal cycling of materials, high conservation of material and no generated waste or escaped heat energy.[31]

The concept of sustainability encompasses development that not only meets the needs of the present but also does not compromise the ability of future generations to meet their own needs. Sustainability by itself is a concept and not a tool. It is a goal toward which businesses use various technological tools to amend their actions. The predominant strategy employed in seeking sustainability is to prevent the occurrence of pollution by using materials, processes, and practices that help in reducing or eliminating pollutants at the source itself, thusly including the reduction in the usage

of hazardous materials. Using proper environmental technology portfolio industries can shift their focus and commitment from controlling and mitigating pollution to actually preventing it.[32]

INDUSTRIAL PROFITABILITY AND SUSTAINABLE DEVELOPMENT

The relationship between environmental and economic performance and the influence of corporate environmental strategy choice on this relationship was studied by Wagner and Schaltegger. After formulating a theoretical model, and performing an empirical analysis focusing on the European manufacturing industry, they found that for firms with shareholder value–oriented strategies, the relationship between environmental performance and different dimensions of economic performance is more positive than for firms without such a strategy.[33]

As presented before, Porter reported that according to detailed case studies of hundreds of industries, based on dozens of countries, internationally competitive companies are those with the capacity to improve and innovate continually. He argues that properly designed environmental standards can trigger innovation that can offset the costs of complying with them. Reducing pollution often coincides with improving productivity. By stimulating innovation, strict environmental regulations can actually enhance competitiveness.[34]

Business can benefit directly from pursuing environmental projects because regulation in areas such as energy efficiency and waste reduction can deliver cost savings and help companies develop more attractive products. These reduced costs add up to substantial benefits across the whole economy. For example, research in the United Kingdom suggests that waste minimization could yield savings of almost 4.4 billion euros in manufacturers' annual operating costs, equal to 7 percent of profits in 2000. Notably, 60 percent of the savings come from the costs of materials that do not end up in the final product. About 2.7 billion euros were saved by industry through energy savings alone. The typical payback periods for waste investments are reported to be about 12 months. Implementation of environmental management systems and practices promise savings of 1.3 billion euros in the agriculture sector.[35] The health care company Baxter International determined that it is saving more than 50 million euros a year upon implementing waste management measures which included changing packaging techniques and new waste reduction strategies. The pollution prevention program at 3M that started in 1975 has saved the company over 740 million euros to date. The results of cost-benefit analysis for recovery and recycling of differentiated paper and cardboard at Italian National Consortium for the Recovery and Recycling of Cellulose Based Packaging (Comieco) shows a positive balance of 610 million euros, the equivalent of the entire yearly production of the Italian paper industry and the equivalent of 3.5 years of paper consumption of the newspaper industry.[36]

King and Lenox showed that profitability factor can be best estimated for disaggregated pollution prevention components. They proposed a link between pollution reduction methods and financial performance and suggested a method through which firms can evaluate whether the relationship is real or an artifact of other firm attributes.[37] They also suggested that in order to maximize the profit, pollution

reduction strategies at a minimum should show that the ratio of the marginal productivity of each activity to the cost of that activity remains the same. As a result, for a profit-maximizing firm, the marginal cost of reducing a unit of pollution will be the same for all pollution reduction options and equal to the marginal benefit of pollution reduction.[38]

Increased process innovation is often associated with unexpected benefit from waste prevention since it could allow for improving measurement of the production process, and thereby, facilitating process innovation.

As reported by King, mandated wastewater pollution control in a printed circuit board industry has led to changes in organizational design, and unexpected and highly profitable process improvement through innovations.[39] He explains how creation of a pollution control department helped to both improve the efficiency of the production and to reduce the pollution at the printed circuit board industry. The main role played by this department was to facilitate the flow of information from different facets of the organization. The new initiative allowed unique access to, and perspective on, information from inside and outside the organization and therefore helped identified incentives, which would have been otherwise overlooked in efforts aimed at improving core industrial processes.[40]

Utilizing data from selected Standard & Poor's (S&P) 500 firms, Hart et al. examined whether pollution abatement imposes costs on firms or is instrumental to better competitiveness by providing cost saving advantages.[41] Their study suggested a direct relationship between environmental best practices and firm performance and showed that a cost advantage can result from adopting best practices in production processes.

Other studies confirm these findings and have shown that practices benefit from redesigning production processes to be less polluting, substituting less polluting inputs, recycling by-products of processes, and innovating less polluting processes. Such practices are intended to reduce the cost of production by increasing the efficiency of production processes and reducing input and waste disposal costs.[42]

The connection between environmental management strategies employed by the firms and firm's financial demographics has been the focus of many recent studies. Reinhardt (1998) showed that not all firms might be able to create competitive advantage from implementing environmentally responsible strategies. Based on his analysis, more attention needs to be paid to the circumstances under which responsible environmental strategies can contribute to the competitiveness. His examples of environmental product differentiation suggest that whether or not a firm can gain differentiation advantage from being environmentally responsible primarily depends on external contingencies, such as the structure of the industry and characteristics of the product market in which a firm competes. Accordingly, resources and capabilities that are developed and used in firms' other productive activities might be required to successfully implement process-focused best practices of environmental management in order to generate all the potentially associated cost savings (complementary assets that are developed in the course of other productive activities, i.e., general business strategies).[43]

It is, however, important to note that the state of the economy can significantly influence company's decision to pursue beyond compliance programs. While controlling cost perceived as a necessity, and where day-to-day compliance has largely been attained, the need for continued spending on beyond compliance seems unsound to

many corporate business executives. Under these conditions, internal support for beyond compliance initiatives may be difficult to muster.[44]

POLLUTION TRADING AND FIRMS FINANCIAL PERFORMANCE

Companies that possess successful environmental programs (environmental management and environmental health and safety) and have been able to achieve regulatory compliance view beyond compliance spending as a fruitless exercise. In most cases, corporate environmental health and safety executives are being asked to rationalize continued investments in their capabilities and staff, or even justify their very existence.

The new pollution abatement initiatives, such as Emission Trading (ET), have impacted industry and firms in a positive manner. By providing flexibility, such innovative programs achieve both the environmental objective and a reduction in compliance costs. An estimate of the cost savings from emissions trading requires both a measure of the amount of emissions trading, and an estimate of the cost of an assumed and/or hypothesized alternative to trading. Though it is difficult to quantify the cost savings to industry via implementing Title IV, ET program according to the available data, analysis to date have shown that those savings are real, and substantial. The relevant literature suggests that upon passage and implementation of The U.S. Acid Rain Program, title IV of the 1990 Clean Air Act Amendments, emissions trading has had a positive outcome for the industry. It has reduced the cost of compliance and impacted industry bottom line favorably. Data also suggests that the reconciliation of the allowances and the emissions, has been significant in providing financial incentives to the industry. For example, in 1995 about 45 percent used grandfather allowances implied cost savings from either spatial or intertemporal emissions trading. Setting aside the actual cost of acquiring allowances (which may have been zero), the opportunity cost estimated for selling allowance in 1995 has been substantial.[45,46]

Firm-level data about electric utilities was used by Considine and Larson to develop an empirical model of how electric utilities use and bank sulfur dioxide (SO_2) pollution permits under the Acid Rain Program. The empirical model considers emissions, fuels, and labor as variable inputs with quasi-fixed stocks of permits and capital. Substitution possibilities between the environment and other production factors examined. The empirical findings indicated that firms bank permits primarily as a hedge against uncertainty and for other firm-specific reasons. Overall, based on their findings, it was suggested that cap-and-trade approaches can reduce the cost of meeting environmental goals by providing a mechanism for addressing regulatory and market risks and by signaling an appropriate price for factor use, especially irreversible capital investments.[47]

The introduction of mandatory controls and a trading scheme covering approximately half of all carbon dioxide emissions across Europe has triggered a debate about the impact of emissions trading on the competitiveness of European industry. Economic theory suggests that, in many sectors, businesses will pass on costs to customers and make net profits due to the impact on product prices combined with the extensive free allocations of allowances.[48]

Robin Smale et al. applied the Cournot representation of an oligopoly market to five energy-intensive sectors (cement, newsprint, steel, aluminum, and petroleum) to investigate the impact of introduction of mandatory controls and trading carbon dioxide (EU ETS) on competitiveness in European industry. The results of this study showed that trading resulted in increase in the cost of the product, extension of the cost to customers, changes in the industrial output, changes in some industries' UK market share, and changes in the firm profits.[49] Their study concluded that The EU ETS delivers emissions reductions and has a positive (or at least non-negative) impact on earnings before interest, tax, depreciation and amortization (EBITDA). Specifically, companies responded to the increase in marginal cost brought about by the EU ETS by cutting back output and increasing prices to cover the additional costs, while simultaneously benefiting from the free allocation of grandfathered allowances. Despite all the changes, their data showed that most participating sectors expected to profit from the trading program although a modest loss of market share was observed for steel and cement, and closure in the case of aluminum.[50]

CONCLUSION

Through innovative design, creation, processing, use, and disposal of substances, industry plays a major role in advancing applications to support sustainability and allow for environmental, economic, and societal growth. Historically, economic and environmental goals have been perceived as conflicting forces with a common belief that economic criteria must be satisfied before environmental goals are pursued. The conception of sustainable development, however, argues that the goals of environmental and economic growth are neither mutually exclusive nor necessarily conflicting. A comprehensive analysis of the environmental risk mitigation strategies through application of environmental technology at selected leading organizations indicated that in addition to traditional criteria, such as cost, quality, and performance, environment is becoming a critical operating component in both product realization processes and operations. It has also been observed that including environmental decision making throughout the entire value adding processes could result in optimally mitigated risks, and compatible economic and environmental growth. Successful integration of environmental values into the operation context, however, should involve both recognition and renovation of a complex mix of interacting factors. Continuous environmental improvement, which is also economically beneficial, can be achieved by enhancing efficiency and productivity of industrial systems, and maintaining material use, recycling, and reuse in the industrial context just as techniques for proficient use of materials and energy are explored.

NOTES

1. Stefan Schaltegger and Terje Synnestvedt, "The Link between 'Green' and Economic Success: Environmental Management as the Crucial Trigger between Environmental and Economic Performance." Available online October 16, 2002. *Journal of Environmental Management* 65(4) (August 2002): 339–346.

2. D. G. Sheldon, "Prioritization of Economic and Environmental Concerns in a Transitional Society," Georgia Institute of Technology, 2000, SD Gen, cherry.gatech.edu (accessed February 2008).
3. Ronald Inglehart, "The American Political Science Review." 75(4): 880–900.
4. World Commission on Environment and Development, *An Overview*. Oxford: Oxford University Press, 1987, www.wsu.edu (accessed February 2008).
5. Stefan Schaltegger and Terje Synnestvedt, "The Link between 'Green' and Economic Success: Environmental Management as the Crucial Trigger between Environmental and Economic Performance," Available online October 16, 2002. *Journal of Environmental Management*, Volume 65, Issue 4, August 2002: 339–346.
6. T. Kronenberg and S. Fuss, Proceedings, SSES Annual Meeting, Zurich, 2005.
7. S. Borghesi, FEEM Working Paper No. 85-99, 1999, http://ssrn.com/abstract=200556 (accessed February 2008).
8. Ibid.
9. G. M. Grossman, A. B. Kruger, "Economic growth and the environment," *Quarterly Journal of Economics* 110 (1995): 353–377.
10. International Society for Ecological Economics, Internet Encyclopedia of Ecological Economics, David I. Stern, Department of Economics, Rensselaer Polytechnic Institute, Troy, New York, 2003.
11. J. William Sugar and Linda Descano, "Identifying the Business Value of Superior Environmental Performance: Current Deliberations from the Aspen Institute." Accessed September 15, 2008.
12. William Baue, U.S. Securities and Exchange Commission, "SEC Urged to Strengthen Rules Governing Corporate Disclosure of Environmental Risks." August 21, 2002.
13. Nicholas E. Costaras, "Environmental Risk Rating for the Financial Sector," *Journal of Cleaner Production* 4(1) (1996): 17–20.
14. M. J. Kennedy, "The Management of Environmental Risk in a Global Industrial Company." *Corporate Environmental Strategy* 8(2) (July 2001): 177–185.
15. P. N. Sharratt and P. M. Choong, "A Life-cycle Framework to Analyze Business Risk in Process Industry Projects." *Journal of Cleaner Production* 10(5) (October 2002): 479–493.
16. S. T. Hart, "A natural resource-based view of the firm." *The Academy of Management Review* 20(4) (1995): 986–1014.
17. A. B. Jaffe, S. R. Peterson, P. R. Portney, and R. N. Stavins, "The Energy Efficiency Gap: What Does It Mean?" *Journal of Economic Literature* 33(1) (1995): 132–163.
18. A. M. Rugman and A. Verbeke, "Corporate strategies and environmental regulations: an organizing framework." *Strategic Management Journal*, *Special Issue: Editor's Choice* 19(4) (1998): 363–375.
19. Ibid.
20. M. E. Porter and C. Van der Linde, "Toward a New Conception of the Environment-Competitiveness Relationship." *The Journal of Economic Perspectives* 9(4) (1995): 97–118.
21. D. A. Rondinelli and A. M. Berry, "Air Pollution in the 21st Century: Priority Issues and Policy Facing the Air Pollution Agenda for the 21st Century," US-Dutch symposium No. 5, *Studies in Environmental Science Journal* 72 (1998): 923–946; ISSN 0166-1116, Congrès PAYS-BAS (SD).
22. B. R. Allenby and D. J. Richards, *The Greening Industrial Ecosystem* (Washington, DC: National Academy of Engineering, National Academy Press, 1994).
23. A. King and M. Lenox, "Exploring the Locus of Profitable Pollution Reduction" *Management Science* 48(2) (2002): 289–299.
24. J. P. Womack, D. T. Jones, and D. Roos, *The Machine that Changed the World*: *The Story of Lean Production* (New York: Harper Perennial, 1991).

25. See note 22.
26. B. R. Allenby, "Implementing Industrial Ecology: The AT&T Matrix System." *Interface* 30 (2000): 42–54.
27. S. Sharma and S. Vredenburg, "Proactive Corporate Environmental Strategy and the Development of Competitively Valuable Organizational Capabilities." *Strategic Management Journal* 19(8) (1998): 729–753.
28. Enrique Claver, María D. López, José F. Molina, and Juan J. Tarí, "Environmental Management and Firm Performance: A Case Study. *Journal of Environmental Management* 84(4) (2007): 606–619.
29. B. R. Allenby and D. J. Richards, *The Greening Industrial Ecosystem,* (Washington, DC: National Academy of Engineering, National Academy Press, 1994).
30. See note 26.
31. See note 29.
32. L. Nilson, "Introduction to Industrial Environmental Management and Cleaner Production, www.ima.kth.se/im/3c1352/text/CPIntroductiontext.pdf (last accessed March 2008).
33. Marcus Wagner and Stefan Schaltegger, "The Effect of Corporate Environmental Strategy Choice and Environmental Performance on Competitiveness and Economic Performance: An Empirical Study of EU Manufacturing." *European Management Journal* 22(5) (October 2004): 557–572.
34. See note 20.
35. "The Contribution of Good Environmental Regulation to Competitiveness." Paper by the Network of Heads of European Environment Protection Agencies, November 2005, www.eea.europa.eu/documents/prague_statement/prague_statement-en.pdf (last accessed March 2008).
36. Ibid.
37. See note 23.
38. Ibid.
39. A. King, "Engineering Management." *IEEE Transactions* 42(3) (1995): 270–277.
40. Ibid.
41. S. L. Hart, A. Gautam, "Does It Pay to Be Green? An Empirical Examination of the Relationship between Emission Reduction and Firm Performance." *Business Strategy and the Environment* 5 (1996): 30–37.
42. P. Christmann, "Effects of 'Best Practices' of Environmental Management on Cost Advantage: The Role of Complementary Assets." *The Academy of Management Journal* 43(4) (2000): 663–680.
43. Ibid.
44. P. A. Soyka, S. J. Feldman, "Capturing the Business Value of EH&S Excellence." *Corporate Environmental Strategy* 5(2) (2001): 61–68.
45. A. B. Jaffe, S. R. Peterson, P. R. Portney, and R. N. Stavins, "Environmental Regulation and the Competitiveness of U.S. Manufacturing: What Does the Evidence Tell Us." *Journal of Economic Literature* 33(1) (1995): 132–163.
46. Denny Ellerman, Richard Schmalensee, Paul L. Joskow, Juan Pablo Montero, and Elizabeth M. Bailey, "Emission Trading under the U.S. Acid Rain Program: Evaluation of Compliance Costs and Allowance Market Performance." Center for Energy and Environmental Policy Research, Massachusetts Institute of Technology, (1997).
47. Timothy J. Considine and Donald F. Larson, "The Environment as a Factor of Production." *Journal of Environmental* 52(3) (November 2006): 645–662.
48. Robin Smale, Murray Hartley, Cameron Hepburn, John Ward, and Michael Grubb, "The Impact of CO_2 Emissions Trading on Firm Profits and Market Prices."*Climate Policy* 6 (2006): 31–48.
49. Ibid., pp. 29–46.
50. Ibid.

BIBLIOGRAPHY

Baue, William. (2002, August 21). U.S. Securities and Exchange Commission "SEC Urged to Strengthen Rules Governing Corporate Disclosure of Environmental Risks."

Christmann, Petra. "Effects of 'Best Practices' of Environmental Management on Cost Advantage: The Role of Complementary Assets." (2000). *Academy of Management Journal* 43(4): 663–680.

Claver, Enrique, María D. López, José F. Molina, and Juan J. Tarí. (2007, September). "Environmental Management and Firm Performance: A Case Study." *Journal of Environmental Management* 84(4): 606–619.

Considine, Timothy J., and Donald F. Larson. (2006, November). "The Environment as a Factor of Production." *Journal of Environmental Economics and Management* 52(3): 645–662.

Costaras, Nicholas E. (1996). "Environmental Risk Rating for the Financial Sector." *Journal of Cleaner Production*" 4(1): 17–20.

Kennedy, M. J. (2001, July). "The Management of Environmental Risk in a Global Industrial Company." *Corporate Environmental Strategy* 8(2): 177–185.

Rondinelli, D. A., and A. M. A. Berry. (1998). "Air Pollution in the 21st Century: Priority Issues and Policy, Facing the Air Pollution Agenda for the 21st Century." U.S.-Dutch symposium No. 5, *Studies in Environmental Science Journal* 72: 923–946.

Schaltegger, Stefan, and Terje Synnestvedt. (2002)."The Link between 'Green' and Economic Success: Environmental Management as the Crucial Trigger between Environmental and Economic Performance," Available online October 16. *Journal of Environmental Management* 65(4) (August): 339–346.

Sharratt, P. N., and P. M. Choong. (2002, October). "A Life-Cycle Framework to Analyze Business Risk in Process Industry Projects." *Journal of Cleaner Production* 10(5): 479–493.

Smale, Robin, Murray Hartley, Cameron Hepburn, John Ward, and Michael Grubb. (2006). "The Impact of CO_2 Emissions Trading on Firm Profits and Market Prices." *Climate Policy* 6: 31–48.

Soyka, P. A., and S. J. Feldman. (1998). "Capturing the Business Value of EH&S Excellence." *Corporate Environmental Strategy* 5(2): 61–68.

Sugar, J. William, and Linda Descano. (2000). "Identifying the Business Value of Superior Environmental Performance: Current Deliberations from the Aspen Institute," available online September 15.

CHAPTER 19

Beyond Segregation of Duties: Next-Generation Techniques in Evaluating User Access Control Risks

Jeffrey T. Hare

INTRODUCTION

Enron, WorldCom, Kanebo, Livedoor, Murakami Fund, Tyco, Adelphia, Peregrine Systems. Household names turned from famous to infamous almost overnight. Scandals rocked these companies and the markets throughout the world. The result of such failures has caused a wave of new and stricter corporate governance rules throughout the world—Sarbanes Oxley (SOX), Japan's Financial Instruments and Exchange Law (JSOX), Basel II, Solvency II, Corporate Law Economic Reform Program (Audit Reform & Corporate Disclosure) Act, and scaled down versions of SOX in both Canada (CSOX) and Europe (EuroSOX) to name a few notable examples.

The revolution in the audit community has been dramatic: new risks to address, new audit procedures to develop, new audit reports, new training for their auditors, new regulations to review, and new and changing auditing standards.

One of implications of the various internal controls audit standards has been greater scrutiny on application security and segregation of duties (SOD), in particular. The interpretation of these new standards by various audit firms, including the Big Four, has required companies to implement controls related to SOD. SOD controls are, after all, a critical element to the prevention of fraud which is, in part, what leads to many of these companies' failures.

In order to design and implement controls related to SOD and application security, firms need to take a risk-based approach. Amazingly enough, even several years after the introduction of many of these internal control audit standards, there is little consensus on the methodology and content needed to perform such a risk assessment.

USER ACCESS CONTROLS, NOT JUST SEGREGATION OF DUTIES

Much of the concern related to application security has focused on segregation of duties. However, as we shall see, there are risks beyond traditional SOD risks

that need to be considered. Further, we will see that the risk assessment process related to SOD conflicts are not yet mature. What, then, will the next-generation risk assessment process look like? We will first look at the elements of a well-designed risk assessment framework. Then, we will address the types of risks that need to be evaluated in this next generation analysis.

For those of you old enough to know of the singing artist Prince, you'll remember that at some point in his illustrious career he became known as "The Artist Formerly Known as Prince." The next generation of segregation of duties will someday have a similar moniker. Whereas today's buzz phrase related to application security design is segregation of duties, as we will see, the risk assessment context needs to be broader, introducing "user access controls . . . formerly known as segregation of duties."

RISK ASSESSMENT METHODOLOGY

What are the components necessary to perform a risk assessment analysis over user access controls? Any methodology must allow for differences in risk tolerances that exist at various companies. Whereas one company's tolerance for fraud may be high because of its size or trust in its employees, it may be very low at others because of past experience. Whereas one company may be willing and able to segregate duties easily because they have adequate staff, another company may want or need to run lean and not have the ability to segregate certain functions. Whereas one company may trust its seasoned employees, another company may recognize that fraud is most often committed by employees with the longest tenure.

Probably the most significant key to success in a risk assessment is having the right personnel involved in the process. The types of personnel that should be included in a risk assessment process are as follows:

- Those who understand the process documents and the process *in practice* in case the "actual" process is different from what is documented.
- Functional managers and staff in the areas being evaluated since some processes at the staff level may happen differently than managers think (i.e., due to poor training, problems with the system, or manual workarounds).
- Staff involved in the testing of the controls on a regular basis (may be management, internal audit, or a separate group like a corporate governance department, depending on the maturity of your process and size of your organization).
- Those who are trained in risk assessment methodology and best practices (risk managers, certified public accountants, internal auditors, and compliance managers).
- Information technology (IT) staff that develops and supports the applications involved in the process.

In its most simplistic overview, following are the steps to a risk assessment process.

First, begin with the most comprehensive "conflict" matrix. As we will see from further discussion, conflict matrices are still evolving. Many of the risk advisory and audit firms have a long way to go in the evolution of their matrices. Ideally, seek an expert in the system or systems you are evaluating. Each system has unique risks,

and it's these risks that are critical to understand in order for the risk assessment process to be mature and complete. One example is Oracle's E-Business Suite, which has a series of forms that allow structured query language (SQL) statements to be embedded and executed within them. These forms are unique to that application and will not be found in SAP or other enterprise resource planning (ERP) systems. However, other ERP systems may have forms with similar risks that need to be evaluated.

Second, start with the most comprehensive definition of risk for each conflict. As we will see from further discussion, this area is still evolving and has been a source of frustration for many companies that were given inadequate or faulty risks during their audit. Ideally, seek a risk advisory firm that is well known and published on this topic. As you are interviewing firms to consider, have them give samples of conflicts and risks their matrix provides.

Third, identify any controls that would mitigate some or all of the risk for each conflict. This is an art, not a science, and this process takes people from various disciplines in order to be successful.

Fourth, assess the residual risk after considering the controls identified. Take into account the operating effectiveness of the control; that is, whether the risk that the control was designed properly, but may not be operating effectively. Therefore, the control may not truly mitigate the risk. Also, consider past testing results and any design deficiencies noted by your internal or external auditors.

Fifth, you may want to consider as a group to prioritize the risks with a residual risk rating. First, agree on a scale. The scale could be something simple like low, medium, and high. Usually, the residual risk rating is subjective. The goal of identifying a residual risk rating would be to come to a consensus as a group. The outcome of identifying a residual risk rating would be that the higher risks would be addressed first.

Finally, management's disposition of the residual risk should be documented. Often, the results of a thorough risk assessment process are changes such as:

- Access is removed from one or more employees.
- Application security is changed.
- Business process is changed.
- Controls are changed or additional controls are put in place.
- Testing cycles are added to certain controls or frequency of testing of the controls is changed.
- Automation of controls is considered (either vendor supplied or customized).
- Implementation of software to monitor or prevent user access control (formerly known as SOD) conflicts.

THE NEXT GENERATION OF SEGREGATION OF DUTIES: USER ACCESS CONTROLS

Even with the right staff, well-planned meetings, and the right methodology, a risk assessment process can be faulty if the conflicts and risks identified are not complete. The goal, then, is to start with the most comprehensive conflict matrix that includes all known risks when starting the risk assessment process.

What is the next generation of segregation of duties? We will look at several components that are necessary in this next generation of segregation-of-duties conflict matrices. The risk assessment process should consider:

- Traditional SOD (i.e., two-function) risks.
- Single function risks.
- Risks related to accessing sensitive data.
- Account processes outside the system as well as that which happens in the system.
- The possibility of submaterial fraud.
- The uniqueness of each company.
- Special risks that privileged users present.

Traditional Segregation of Duties

A comprehensive risk assessment process takes into account traditional segregation of duties risks. As it relates to access controls, companies first think of segregation of duties (SOD). SOD encompasses the risk of a user having access to two functions that taken together allow a certain risk. For example, a user having both the ability to enter a supplier and enter an invoice related to that supplier may have the ability to commit fraud, absent any controls that would detect the fraudulent entries. The most critical element to a company's success is in the understanding and assessing risks related to each SOD conflict. When assessing risk, as we will later discuss in more detail, unless the person or group performing the risk assessment truly understands the risk, they will not know whether or not the controls they identify to mitigate such risk truly have the mitigating impact desired.

Let's look at the example of "enter supplier versus enter invoice" in more detail to better understand this concept. The primary risk of access by the same person having both functions is that they could commit fraud. They could start by entering a fictitious supplier with a personal post office box or mailing address that directs the check to them or an accomplice. Next, they enter an invoice against such a supplier. There are many possible controls you could identify that could help keep such a fraud from being committed. The control(s) that you would identify would be known as a mitigating control because such a control would help to mitigate the risk of the fraud. Let's look at some examples of ways this type of fraud could be "caught":

- The person signing the checks could review the supporting documentation and question the nature and/or validity of the invoice or question the validity of the vendor.
- If checks are automatically signed, perhaps the controller review would review a payment register and review the supporting documentation before the checks are released.
- If the invoice were coded to an account that receives a lot of scrutiny such as travel expense, the department manager in charge of that budget could uncover the fraud as they look at the details of the expenses for that period.

What happens, however, if those performing the risk assessment process are focused on some risks but not others? For example, if the group performing the risk

assessment were primarily focused on the risk of material misstatement rather than submaterial fraud, they may feel comfortable relying on controls that only partially mitigate the risk.

Case Study

An international manufacturing company based in the United States had one of the Big Four firms help them design and implement SOD controls primarily in response to requirements under the U.S. Sarbanes-Oxley Act (SOX). In the process of evaluating risks related to the SOD conflict—enter suppliers versus enter AP invoices—they identified a mitigating control that the controller reviewed a payment register for all check runs and looked at all the detail supporting checks over $25,000 prior to the checks' being released.

This control was a reasonable mitigating control to identify given the goal of the control was to prevent material misstatements in their financial statements. However, the risk of material fraud is not the only risk. There is also a risk of submaterial fraud (fraud that would not rise to the level of materiality as it relates to a financial statement audit). The company and Big Four firm that designed these controls failed to take into account the potential for fraud below the $25,000 level.

After performing a risk assessment on this particular conflict and in response to the submaterial fraud risk, the company added a sampling of payments below the $25,000 level. Management felt the sample of checks below $25,000 was a reasonable (not absolute) deterrent to someone considering committing fraud.

In this case study, the risk assessment process failed to take into account certain risks and left their company vulnerable to fraud below the level of review by the controller. However, after understanding all risks posed with that conflict, management made a decision to change the review process.

Single-Function Risks

A comprehensive risk assessment process takes into account single-function risks. The biggest gap in the way that many of the audit and risk advisory firms have approached this topic is that they have typically viewed access control risks only in the traditional SOD paradigm. That is, they have identified "conflicts" as always being between two functions. Our example involves the conflict between entering suppliers and entering an invoice against such a supplier. However, there are cases where the real risk lies in having access to a single function.

Let's look at an example. One of the most significant fraud risks relates to the maintenance of bank account information for suppliers being paid via an automated clearinghouse (ACH). Absent any mitigating controls, an employee with access to this function can change a bank account the day before a payment run for a high-volume supplier to redirect the payment to their bank account.

Exhibit 19.1 is an example of some conflicts identified from one of the big audit firms.

The risk noted in each of these examples is the risk of "payments to inappropriate bank accounts." This is a risk related to having access to update the bank account information. However, a fraudster doesn't need access to the Requisition Templates, Returns, or Update Accounting Entries function in order to commit fraud.

EXHIBIT 19.1 Sample Conflicts in Segregation of Duties

Process 1	Process 2	Risk Noted
Banks	Requisition templates	Payments to inappropriate bank accounts
Banks	Returns	Payments to inappropriate bank accounts
Banks	Update accounting entries	Payments to inappropriate bank accounts

As you can see from this example, this particular audit firm isn't really taking a risk-based approach to developing their conflict matrix. The risk noted for each of these conflicts is the same and the second process (Requisition Templates, Returns, Update Accounting Entries) isn't needed in order to commit fraud. In this case, the bank's function is a high-risk, sensitive function and, therefore, should be evaluated on its own as a single function risk.

Exhibit 19.2 shows some examples of other single functions with high risks.

Sensitive Data

A comprehensive risk assessment process takes into account access to sensitive data. Data theft is becoming big business throughout the world of organized crime. Any analysis of user access controls should include access to sensitive data. The first challenge is determining what constitutes sensitive data. There are three categories of data that should be considered: regulatory, proprietary, and other sensitive data.

First, you need to identify the types of data that needs to be considered due to regulatory requirements. There are some categories of data that are subject to industry regulations such as health insurance–related data (e.g., Health Insurance Portability and Accountability Act [HIPAA])[1] and financial services (e.g., Gramm-Leach-Bliley Act [GLBA])[2] in the United States. There are other categories of data security required for all organizations such as personally identifiable data via regulations, such as state notification laws[3] in the United States and national requirements in Europe.

EXHIBIT 19.2 Sample Single Function Risks in Segregation of Duties

Function	Risk
Remit to addresses	Directing customer remittances to a fictitious PO box.
Suppliers	Setting up of a fictitious supplier. Depending on other controls, could commit fraud by mailing in an invoice that doesn't go through an approval process such as utilities or rents.
Role definition	Changing of security access to grant access to sensitive data or a function that gives them the ability to manipulate data and/or commit fraud.
SQL Forms	Embedding an SQL statement that updates data with fictitious information, resets the password for a powerful log-in such as one with system administrator privileges, or allows circumvents the change management process.
Purchasing locations	Directs inventory to a fictitious warehouse allowing theft of inventory.

Second, there is some data that is proprietary to a company. Examples would include manufacturing formulas, suppliers, and customers. This data would be important to keep out of the hands of its competitors.

Finally, there is data that may not be covered by regulatory requirements or proprietary to the company, but should still be protected. Examples would include supplier and customer bank accounts. Protecting such data may not be required, but would be considered a good business practice.

Takes into Account Processes Outside the System and Those in the System

A comprehensive risk assessment process takes into account process outside the system as well as what happens within the system. While there are certain risks of fraud within the IT system, any assessment of fraud includes an understanding that some fraud happens outside system or part of the process happens outside the system. There are many fraud schemes that start with the theft of an asset and then the fraudster attempts to cover up the theft using their normal duties (within or outside the IT system) or in collusion with an accomplice such as a supplier, customer, or another employee.

A well-designed risk assessment process would include "conflicts" such as those shown in Exhibit 19.3.

Consider the Possibility of Submaterial Fraud

A comprehensive risk assessment process takes into account the risk of submaterial fraud. As we saw in our case study noted earlier, it is possible to design controls to effectively catch material misstatements in the company's financial statements, yet which fail to mitigate risks below the materiality level. In the earlier example, the procure-to-pay process included a detailed review of all checks greater than $25,000. However, there were no controls implemented to prevent or detect fraud below the $25,000 level.

EXHIBIT 19.3 Sample Segregation-of-Duties Risk Assessment

Process	Conflicting Process	Risk
Access to income cash from customers	Writing off transactions or balances	Theft of incoming cash and concealment by writing off specific transactions or balances for that customer.
Access to check stock	Reconciliation of bank statement	Writing of a fictitious check and concealment of such during bank reconciliation process by requesting a journal entry (include with bank charges) or burying in a reconciling line item.
Requesting or approving a new supplier	Entering a purchase order	Establishing a fictitious supplier and issuing a PO (two-way match) against such a supplier. Then, mailing in an invoice that references the fictitious PO.

Throughout the world, many of the firms involved in the risk advisory process are also external auditors or were formed by those with an external audit background. Whereas fraud examiners and internal auditors have been concerned about submaterial fraud, financial statement auditors have been concerned about fraud that could have an impact on the financial statements (i.e., material fraud). That is not to say that submaterial fraud has never been the concern of external auditors. If external auditors discover fraud during the internal controls or financial statement audit, they would need to determine if that fraud could rise to the level of a material misstatement. Such an assessment would include considering that such fraud could have been ongoing for several years and the impact of the fraud of the reputation of company. However, because of the deficiencies in the content and methodology of their fraud assessment, fraud below the materiality level could go undetected by external auditors.

Management needs to recognize that the prevention and detection of submaterial fraud is an important risk to address. Therefore, the risk assessment process needs to take into account such risks.[4]

Takes into Account the Uniqueness of Each Company

A comprehensive risk assessment practice takes into account the uniqueness of each company. Each company is not only unique in its business, but is also unique in its process design, controls, risk tolerance, and IT systems. Each of these elements adds something to the risk assessment process.

First, each company has a unique process design. The order to cash cycle is probably the best example of this because of the variation in revenue cycles. There are significant differences between an online retailer, whose primary business is credit card transactions over the Internet, and a software company. The online retailer is not concerned about creditworthiness or revenue recognition, but is very concerned about protected credit card information and complying with the provisions of the payment card industry, whereas the software company would be very concerned about the creditworthiness and revenue recognition issues, but would likely not have exposure to credit card issues.

Second, each company has a unique controls design. Similar to differences in business process design, you'd also expect significant variation in the design of internal controls from company to company. You wouldn't expect an online retailer to have significant controls related to revenue recognition, but you would expect a software company to have several controls and several levels of review related to revenue recognition. Alternatively, you would not expect a software company that doesn't process credit cards to have controls over the storage and access to credit card data, but you would have those expectations of the online retailer.

Third, risk tolerances vary from company to company. A company's risk tolerance is often a product of its senior management and board of directors. Their risk tolerance often is a function of their experience and background as well as their advisers. In our procure-to-pay case study, where one chief financial officer may be willing to assume the risk of fraud below the $25,000 level, another may not because of experience with similar fraud at that or another company.

Finally, IT systems vary significantly and differences need to be considered. If you are evaluating the risk of someone committing fraud in the procure-to-pay process,

the variations in processing in the payables module may influence the risk assessment process. If a system allows for the processing of a payment without the entry of an invoice, then a user having the ability to enter a supplier and process a payment is high-risk conflict. However, if the IT system doesn't allow a payment to be made without the entry of an invoice, then the risk shifts to a user having the ability to enter a supplier and enter an invoice. The risk of a user having the ability to enter a supplier and process a payment is minimal.

Considers Privileged Users Such as Those that Support ERP Systems

Finally, a comprehensive risk assessment process takes into account access being granted to privileged users that support the application. In many large ERP systems, the support model allows privileged users to regularly have "super user" privileges in the production environment in order to support the application. Because ERP systems don't have an "audit all" function similar to many mainframe systems, a complete audit trail of the activity for these privileged users cannot be produced from the system.[5] Therefore, the risk assessment process needs to take into account the access privileged users may have access to sensitive data, SOD conflicts, and high-risk single functions, albeit possibly at different times.

CURRENT STATE AND FUTURE DIRECTION OF RISK ADVISORY AND AUDIT FIRMS

Given the above are the characteristics of a comprehensive risk assessment process, what is the current state of risk advisory and audit firms and what future direction do they need to take?

First, let's look at some examples from different audit firms' conflict matrices.

Example 1: Suppliers

Exhibit 19.4 shows several conflicts identified by one audit firm. They took a high-risk function, Suppliers (entry of suppliers), and paired it with several other functions. However, the risk noted does not really reflect the true risk. For example, how could a combination of suppliers and tax certificates, payment terms, or tax groups lead to an "inappropriate payment?" Or how could a combination of suppliers and run-mass-cancel or requisition templates lead to payments to fictitious vendors?

As you can see from this example, the risks identified really are not indicative of the access being allowed by the two processes.

Example 2: Banks

Exhibit 19.5 is another high-risk single-function, bank account entry. From a fraud perspective, access to maintaining bank accounts could lead to fraud by allowing an employee to change the bank account for a supplier being paid via ACH. However, a fraudster doesn't need access to the Requisition Templates, Returns, or Update Accounting Entries function in order to commit fraud.

EXHIBIT 19.4 Sample Conflicts Identified by an Audit Firm

Process 1	Process 2	Risk Noted
Suppliers	Card profiles	Inappropriate payment of expenses
Suppliers	Card programs	Inappropriate payment of expenses
Suppliers	Code sets	Inappropriate payment of expenses
Suppliers	Credit cards	Inappropriate payment of expenses
Suppliers	Expense reports	Inappropriate payment of expenses
Suppliers	GL account sets	Inappropriate payment of expenses
Suppliers	AP accounting periods	Inappropriate payments
Suppliers	Banks	Inappropriate payments
Suppliers	Buyers	Inappropriate payments
Suppliers	Define bank charges	Inappropriate payments
Suppliers	Maintain purchase orders	Inappropriate payments
Suppliers	Match unordered receipts	Inappropriate payments
Suppliers	Open interface invoices	Inappropriate payments
Suppliers	Payment terms	Inappropriate payments
Suppliers	Returns	Inappropriate payments
Suppliers	Signing limits	Inappropriate payments
Suppliers	Supplier item catalog	Inappropriate payments
Suppliers	Supplier lists	Inappropriate payments
Suppliers	Tax certificates	Inappropriate payments
Suppliers	Tax groups	Inappropriate payments
Suppliers	Update accounting entries	Inappropriate payments
Suppliers	Payment batch sets	Inappropriate processing of payments
Suppliers	Control purchasing periods	Payments in the wrong period
Suppliers	Expense account rule	Payments to fictitious vendors
Suppliers	Invoice batches	Payments to fictitious vendors
Suppliers	Invoices	Payments to fictitious vendors
Suppliers	Merge suppliers	Payments to fictitious vendors
Suppliers	Payment batches	Payments to fictitious vendors
Suppliers	Payments	Payments to fictitious vendors
Suppliers	Recurring invoices	Payments to fictitious vendors
Suppliers	Requisition templates	Payments to fictitious vendors
Suppliers	Run-mass-cancel	Payments to fictitious vendors

As you can see from this example, this particular audit firm isn't really taking a risk-based approach to developing their conflict matrix. The risk noted for each of these conflicts is the same, and the second process (Requisition Templates, Returns, Update Accounting Entries) doesn't appear to add to the risk. In this case, the bank's function is a high-risk single function. Risks and access to this function should be evaluated on its own.

EXHIBIT 19.5 Risks Associated with Bank Account Entry

Process 1	Process 2	Risk Noted
Banks	Requisition templates	Payments to inappropriate bank accounts
Banks	Returns	Payments to inappropriate bank accounts
Banks	Update accounting entries	Payments to inappropriate bank accounts

EXHIBIT 19.6 Sample Credit Memo Risks

Process 1	Process 2	Risk Noted
Enter credit memo	Enter customer	Access to "enter credit memo" and "customers quick" will allow a user to potentially create themselves as a customer and then issue a credit against that customer in hopes of receiving a credit payment. Can lead to an overstatement and understatement of revenues and receivables.
Enter credit memo	Maintain customers	Access to "enter credit memo" and "enter customer" will allow a user to potentially create a fictitious customer and create a credit memo for that particular fictitious customer in an attempt for payment. Can cause credit memos to be inappropriately processed and therefore affecting a company's revenue.
Book sales orders	Enter credit memo	Access to "book order and transaction batches" will allow a user to inappropriately create customer orders and then create any sort of accounts receivable (AR) transaction against that order such as an invoice, credit memo, debit memo, etc. Can lead to an understatement of receivables and cash.
Enter AR invoices	Enter credit memo	Access to "enter invoice and create credit memo" will allow a user to create a fictitious invoice and then issue a credit memo for that customer in an attempt for payment to that customer. Can lead to an understatement of revenue and cash.

Example 3: Missing Processes

To illustrate the lack of focus on submaterial fraud risk, let me give you an example of what is missing. The ability to generate a credit memo is one way a fraudster could hide the theft of incoming cash (or theft of a check) from a customer. Exhibit 19.6 shows the only major risks noted with having access to enter credit memos from several well-known audit firms. Not one of them mentions theft of cash in their risks.

Examples of other missing components of many conflict matrices are processes such as access to cash, ability to initiate a wire transfer or sign checks, and account reconciliations.

These illustrations demonstrate another challenge that large risk advisory firms have in their methodology. Whereas the term *integrated audit* refers to the audit of internal controls as part of both the financial statement audit and the audit of internal controls, there is another type of integrated audit necessary. Often, the design and evaluation of business processes outside the system are evaluated by a separate group than the IT auditors that evaluate IT controls such as application security and segregation of duties. As we have seen from several examples, certain risks transcend manual and system process. For example, if the supplier establishment process outside the system allows a buyer to create a supplier and that buyer can also issue a purchase order, then they have a decent opportunity to commit fraud. Another example would be a collector/credit analyst asking a customer to send

a check to them directly. If the mailroom controls would allow that check to be delivered to the analyst without being logged and that analyst also has the ability the request a credit memo or write off a transaction or balance, then that analyst has a reasonable opportunity to commit fraud. Only an "integrated audit" that looks at the risk in the process from manual processes to application security within the system would likely identify such fraud risks. Audit and risk advisory firms need to evolve their content and methodology to address these new paradigms.

CURRENT STATE AND FUTURE DIRECTION OF ERP SOFTWARE VENDORS

Software vendors that provide large ERP systems need to evolve their applications to allow security development, monitoring, and prevention for all user access control risks.

This includes several areas such as:

- Providing necessary granularity in the objects embedded in the application.
- Providing necessary automated controls in its core applications.
- Embedding the necessary technologies in its application licenses.

Providing Necessary Granularity in the Objects Embedded in the Application

Many of the tools that monitor user access control risks (not just SOD risks—remember Prince) look at the objects at the database level to evaluate whether the certain users have conflicts. For example, with the two largest ERP application providers, the tool could be looking for a user with a combination of functions in Oracle's E-Business Suite or a combination of transaction codes (T-codes) in SAP. In order for the user access control monitoring tools to monitor or prevent the risks, the application has to contain the necessary granularity. In many cases, these objects are embedded in the application and, therefore, the application design may need to change.

For example, one of the traditional SOD conflicts that needs to be evaluated would segregate the entry of an order versus the approval of an order. This would mean that one user would be able to enter an order but not approve the order. Another user would be able to view the order and not change it but be able to approve it if the order terms, pricing, and so on are appropriate. In order to monitor such access, the application would need an object that can enter but not approve an order, and an object that can approve but not update an order. Absent appropriate objects, user access control software cannot monitor this conflict. Certainly, if these objects cannot be monitored, access cannot be prevented and, therefore, the controls related to user provisioning cannot be fully automated.

Providing Necessary Automated Controls in Its Core Applications

ERP application vendors also need to continue to evolve their applications to provide more automated controls. Automated controls are necessary for organizations to

reduce risk and thus, audit fees where such automation supports important controls. Use of technologies such as workflow and intelligent design of forms are necessary to meet the needs of organizations.

In the SOD example, an automated workflow process that routes entered orders to those authorized to review and approve the orders would be an ideal automated control. However, the flexibility in designing such a workflow to accommodate the various intricacies of such a process, yet maintaining the integrity of the "enter versus approve order" SOD conflict, is a challenge.

Continued evolution of application design to support these requirements will undoubtedly allow one ERP application vendor to stand out from their competitors.

Embedding the Necessary Technologies in Its Application Licenses

Hundreds of new software firms have been formed over the past several years to address compliance needs. Many of these firms have been bought by the large ERP vendors to address holes in their own application or technology offerings. The expectations of organizations purchasing ERP systems are that elements necessary for compliance initiatives are a part of the application licenses. However, as ERP vendors purchase companies, many of these new applications or technologies are sold as a separate license. Sophisticated buyers of ERP systems are learning more and more that the initial license cost is just the tip of the iceberg when considering all that is necessary to implement the applications in a compliant manner. ERP software vendors will not only need to integrate the purchased applications into their core applications, but they will also need to integrate the licensing of them into the core application license costs.

CONCLUSION

Just as the various internal control audit standards have evolved, the concept of segregation of duties needs to evolve to include all user access control risks. This includes analyzing additional risks such as single function risks and access to sensitive data. It also includes looking at business processes holistically from manual processes to parts of the process that happens inside the IT system.

Risk assessment processes need to evolve to take into account all risks, including submaterial fraud and risks presented by privileged users. The risk assessment process also needs to include the appropriate personnel in order for the process to be comprehensive and effective.

Software firms need to continue to evolve their applications and objects to be able to monitor and prevent all user access controls risks and to allow automation of key processes.

The content and the methodologies used by organizations, audit firms, and risk-advisory firms need to continue to evolve to meet the challenges presented by the next generation of segregation of duties.

NOTES

1. www.hhs.gov/ocr/privacysummary.pdf.
2. www.ftc.gov/privacy/privacyinitiatives/glbact.html.
3. www.ncsl.org/programs/lis/cip/priv/breachlaws.htm.
4. "Sub-Material Fraud Risk: The Elephant in the Room" (white paper), available at www.oubpb.com.
5. "Monitoring Privileged Users in an Oracle Applications Environment" (white paper), available at www.oubpb.com.

CHAPTER 20

Transaction-Based Cross-Enterprise Risk Management

Allan D. Grody and Peter J. Hughes

OVERVIEW

As with all modern enterprises, financial institutions must respond to a myriad of regulations governing finance, the environment, and risk. It is in this latter category "risk" that the challenges for financial institutions are particularly prevalent. The advancing sophistication of financial products and the markets where they are traded have combined with technological innovation to produce a new reality. Financial institutions must now come to terms with the fact that when trades and transactions enter their operating environments they trigger risk exposures that can go well beyond nominal transaction values.

The current financial crisis can be linked to such exponential risk exposures that escalated to billions of dollars without always finding expression in conventional financial accounting and risk reporting systems. The Société Générale fraud and subprime failures are examples of exceptional and unplanned accumulations of risk exposures that escaped the exercising of business judgment simply because executive management was unaware of their existence on such a scale.

What is also evident is that such risk concentrations are not attributable to any particular category of risk. The unreported risk concentrations that contributed to the current financial crisis were a cocktail of all the principal categories of risk—credit, market, liquidity, and operational. Regulators, practitioners, and other thought leaders would be well advised to now focus on the last leg in Basel II's mandate, that of operational risk, in its entirety.

The Bank for International Settlements (BIS) established the Basel Committee on Banking Supervision (BCBS) in 1974, in the aftermath of serious disturbances in international currency and banking markets and deterioration in capital ratios. In 2004, a more demanding Basel Capital Accord ("Basel II") was introduced with three pillars: Pillar One (capital calculation), Pillar Two (regulatory oversight), and Pillar Three (market disclosure). Also new with Basel II is the requirement for operational risk management. While designed for banking, Basel II offers an approach and framework to risk management that all industries may wish to consider.

Basel II classifies operational loss events as resulting from: internal fraud, external fraud, employment practices and workplace safety, clients/products/business

practices, damage to physical assets, business interruption and systems failures, and execution/delivery/process management.[1] In fact, the loss-event categories of operational risk can now be put in perspective and, with hindsight, classified as transaction-based and cross-enterprise risk. Here, many recent events can be attributed in part, if not in total, to problems within these prescribed loss events.

Citibank reported, for example, that its market value at risk (VaR) number did not include collateralized debt obligation positions because they are hard to value in an absence of prices or model inputs[2]; Credit Suisse took a $2.8 billion write-down for valuation-model pricing errors and use of stale prices[3]; Société Générale reported a $4.9 billion loss from trader fraud where improper counterparty codes were used and no systematic ability existed to look across proprietary systems' position data and external exchange position data[4]; and Bear Stearns nearly collapsed because it could not price its mortgage portfolios. This serves to heighten the awareness of financial institutions and their regulators to the need for the measurement and management of risk exposures in the aggregate rather than on a specific risk category or on a "silo" basis.

This silo awareness is expressed in an April 2008 paper issued by the Basel Committee on Banking Supervision entitled "Cross-Sectoral Review of Group-wide Identification and Management of Risk Concentrations." In its introduction the paper explains its aim:

> *". . . to explore the progress that financial conglomerates have made in identifying, measuring, and managing risk concentrations on a firm-wide basis and across the major risks to which the firm is exposed." In commenting on traditional risk management approaches the paper states, "The risk management at financial conglomerates tends to be structured in silos according to the risk category . . . several groups expressed a desire to develop more "horizontal" (i.e., across the risk categories) insight into potential risk concentrations and have started developing management tools to acquire* a more integrated group-wide view of risk exposures and potential risk concentrations."[5]

BACKGROUND

After a decade of intense industry-wide consultation and investment, the operational risk component of Basel II is falling well short of what its founding fathers envisaged when, in its first consultation paper in 1999, the Basel Committee challenged the industry to quantify the level of operational risks and incorporate them into a firm's overall capital adequacy.[6]

A fairly reliable indicator of where the industry is positioned relative to Basel II was the industry's response to the regulatory agencies' implementation plans. Such a process was conducted in 2007 in the United States relative to the joint agencies' request for comment relative to their proposed Basel II supervisory guidance. Consider this response from the Advanced Measurement Approach Group of the Risk Management Association, which was formed to represent the leading U.S. banks in this area:

> *Practically speaking, the requirement to produce comprehensive management reports including "changes in factors signaling an increased risk of future losses" cannot be met at this point in time or in the near future. In*

> *many instances, operational risk factors that led to a particular event cannot be uniquely determined retrospectively, let alone detecting a change in factors that signals an increase in future losses.*[7]

This statement begs the question, what value does a global risk management, capital adequacy, or economic capital regime have if the banks applying it, by their own admission, are unable *at this point in time or in the near future* to fulfill a requirement as fundamental as being able to demonstrate the link between changes in risk factors and past and likely future negative outcomes? The answer is, without this capability, risk management programs and their risk-adjusted performance measures have very little value.

But it is not difficult to see why this set of circumstances exists. An essential ingredient for any risk management program is missing. And that ingredient is "exposure." The industry has not yet found a way of identifying a financial institution's total portfolio of operational exposures in live operating environments and how to put a consistent and comparable value on them.

In the absence of such a direct exposure measurement method the industry has looked to loss history as being the only objective source of information on operational risks. Consequently, the advanced measurement approaches (AMAs) under Basel II rely mainly on loss history to "deduce" the possible current portfolio of operational risk exposures through the application of actuarial modeling techniques.

It is clear by now that a risk management regime that operates on imperfect operational loss history can benefit from a bottom up approach that measures operational risk exposure as well as isolates its root causes. Operational risk exposures fluctuate on a daily basis, often dramatically, as a consequence of changes in transaction volumes, implementations of new technology, failures of existing technology, business reorganizations, staff absences, new products ... the list is endless. There are also hidden exposures related to, for example, fraud and control breakdowns. And if loss events do occur, technology and operations personnel invariably diagnose the causes and fix them.

The conclusion is that historical loss experience is a useful tool to help focus on current exposure but without the ability to measure current exposure we cannot be proactive in risk management or risk mitigation.

BASEL II AND CURRENT U.S. IMPLEMENTATION

The U.S. implementation of the final regulatory guidelines for Basel II related to operational risk calls for a

> *Consistent and comprehensive capture and assessment of data elements needed to identify, measure, monitor, and control the bank's operational risk exposure. This includes identifying the nature, type(s), and underlying cause(s) of the operational loss event(s).*[8]

Basel II states, with respect to understanding and approving the bank's tolerance for operational risk, that:

> *Banks use several approaches to define operational risk tolerance, including establishing expectations for control self assessments, establishing targeted*

> *ceilings for operational losses, developing key risk indicators, or establishing other qualitative expectations for operational risk management. These approaches will continue to evolve and banks are encouraged to develop effective metrics to define their operational risk tolerance.*[9]

Unfortunately, we have not yet achieved a meaningful calibration of operational risk capital nor have we engaged in comprehensive debate on how to measure operational risk. Specifically, a primary reason for failing to arrive at a reasonably useful measure of operational risk is that we have not yet defined the fundamental nature of the measurement unit (or units) of operational risk. We have for all practical purposes deliberately postponed its measurement by defining it in terms of a "qualitative" *assessment process* rather than a "quantitative" *measurement process*. This has left financial institutions to ponder how to link operational risk exposure to their frequency and severity measures of operational losses. If available (and not much is yet available) then operational risk loss data is rather inelegantly utilized to determine the parameters of a typically poorly articulated operational risk model for calculating the 99.9 percent confidence interval over a one-year horizon.

A mapping of loss events into business lines and event types is well on its way in the largest, most internationally active financial institutions that are mandated to comply with the Basel II AMA operational risk approach. Nevertheless, missing from the typical mapping are the causal events at a sufficient level of granularity that resulted in the losses. This failure makes it more difficult to observe risk exposure and perform risk mitigation. Unlike market risk and credit risk, increasing operational risk has no upside, and therefore, every operational loss event is a drain on capital, rather than a calibrated risk for a potential reward.

Accounting for operational risk exposure, and accommodating the operational risk component of risk capital, has proven a formidable challenge and has yet to be structured with anything approaching an accepted model or methodology or even a thoughtful enduring approach. At its most fundamental level, financial institutions have been evolving management information systems over decades. What stands in the way of incorporating risk measures into this management reporting structure is: (1) the lack of any measure of operational risk exposure; (2) the failure to incorporate the importance of data into risk measurement models and, finally; (3) the lack of any cohesive mechanism to correlate operational risk exposure to historical operational losses.

A first step to calculating a risk based operational capital charge calls for understanding the causal events, measuring the risk exposure inherent in the operations associated with those events, and doing so around a common risk measurement framework. We have, unfortunately, failed to develop effective risk metrics in our rush to satisfy the regulators' well-intentioned interest in calculating operational risk capital. We, therefore, begin our quest to resolve this conundrum by first reviewing the current state of risk management and then introducing the concept of transaction based risk accounting.

CURRENT STATE OF ENTERPRISE RISK MANAGEMENT

In order to understand the ultimate aim of enterprise risk management we must first explain the concept of economic capital. Conceptually, economic capital can be

expressed as protection against unexpected future losses and is commonly referred to as the enterprise value at risk (VaR) at a specific confidence level over a particular time horizon. Economic capital is distinct from familiar accounting and regulatory capital measures, and distinct from measures of capital adequacy as mandated under Basel II.

Economic capital is based on a probabilistic assessment of potential future losses and is therefore a potentially more forward-looking measure of capital adequacy than traditional accounting measures.[10] Expressed as the monetary value of capital necessary to adequately support specific risks assumed, most traditional measures of capital adequacy relate existing capital levels to assets or some form of adjusted assets. Economic capital relates capital to risks, regardless of the existence of assets. In the U.S. Basel II has given large financial institutions and their boards the impetus to further develop VaR models through the inclusion of AMAs for market and credit risk (financial risk) and operational risk (nonfinancial risk) and through the application of the basic indicator approach (BIA) for smaller institutions.[11]

In reality, we want to ensure that each business unit incurs an economic capital charge that will allow firms and individual business units to use risk/reward analysis to improve and effectively communicate their operational decisions. A significant challenge for practitioners and academic researchers is to provide the models which will enable a financial institution to calculate the economic operational risk capital saved due to such innovations as internal process improvements, information technology enhancements, impact of external payment and settlement time compression, and so on.

A robust calculation of economic capital depends on being able to properly accumulate historical loss data over time, starting with the accumulation of historical market prices and credit default histories along with adding the accumulated measures of operational risk. Further, it requires the connections between the identities of issuers of debt and equity, and their identities as counterparties in a trade, or as borrowers, to be done in a consistent manner so that risk correlations can be calculated. For example, it is necessary to link a potential defaulting obligor (whose market price of its public debt is declining, a market risk) with its subsidiary that has a loan outstanding whose probability of default is increasing (a credit risk). It is thus apparent that while market risk and credit risk are linked, the diversification benefits between these two financial risks may be affected by added operational risk. This operational risk manifests itself in such granular operational elements as nonstandard counterparty identifiers, poorly articulated hierarchies of business entities, inaccurate cross-references between issuers and obligors, and all manner of manual and automated process that interact with such data elements.

The use of economic capital allows an organization to make objective "risk-adjusted" judgments across business units, including fee-based services, trading desks, credit and deposit businesses, and fiduciary units. A key decision is the amount of capital to allocate to each business line based on its riskiness. In order to manage in this way, the organization must be structured into the appropriate business units and within a hierarchy of accountability in the manner that management information systems require.

Implementing transfer pricing schemes as well as cost accounting and attribution systems are a prerequisite to having a well defined risk based performance management system. Closely aligned with management's responsibility for performance are incentive compensation schemes. These, too, must be in place if performance and

$$\textbf{Risk-adjusted return on capital} = \frac{\textbf{Risk-adjusted revenue}}{\textbf{Economic capital}}$$

Risk-adjusted revenue = (Operating revenue + Return on economic capital)
– (Operating expenses +/– transfer prices – expected losses)

Economic capital = VaR where VaR equals the fully diversified sum of
Credit VaR + Market VaR + Operational risk VaR

EXHIBIT 20.1 Risk-Adjusted Performance Measurement

incentives on a "risk-adjusted" basis are to make sense to the organization (see Exhibit 20.1).

Risk can be divided into losses that are both expected and unexpected. There will be a "normal" amount of loss that business is willing to absorb as a cost of doing business, such as error corrections, frauds, and so on. These failures are explicitly or implicitly budgeted for in the annual business plan and are incorporated into the pricing of the product or service. While we had assumed that a business unit's management was already assessing and pricing expected failures into severe but not catastrophic losses we are now aware that this is not the case. The current financial meltdown has shone a light on management's inability to associate the "riskiness" of each transaction within its own internal operational processes, let alone to external events. And it has exposed the inadequacies of silo structures that characterize large financial institutions wherein management is unable to correlate the riskiness of transactions that occur in one business unit to another and within one category of risk to another. The focus of risk assessment is typically on *unexpected* failures, and the amount of economic capital that should be attributed to business units to absorb these losses.

Exhibit 20.2 illustrates the components of a risk capital calculation whereby:

- *Expected losses* are the anticipated average loss over a defined period of time that represents a cost of doing business and is generally expected to be absorbed by operating income. Expected losses are supposed to be priced into the products' costs and profit margins as, for example, in the case of loan losses where the expected loss is priced into the yield and an appropriate charge included in the reserves provisioned for loan losses. Provisions for credit card losses, payments and securities settlement losses, uncollectible commercial loans, trade counterparty defaults, and so on are estimated as the potential cost of doing business.
- *Unexpected losses* are actual (economic) losses that exceed expected losses and are a measure of the uncertainty inherent in the loss estimate. It is this possibility to incur unexpected losses that necessitates the holding of capital. However, as we have seen, the "expected" failures turn into unexpected losses, and unexpected failures can themselves be further subdivided.
- *Catastrophic losses* are potential risks of losses that can be protected by either the capital of the enterprise, or by insurance, or by mutual risk sharing as in reinsurance, and/or through risk mitigating infrastructure utilities, such as payment networks, settlement systems, and centralized counterparties.

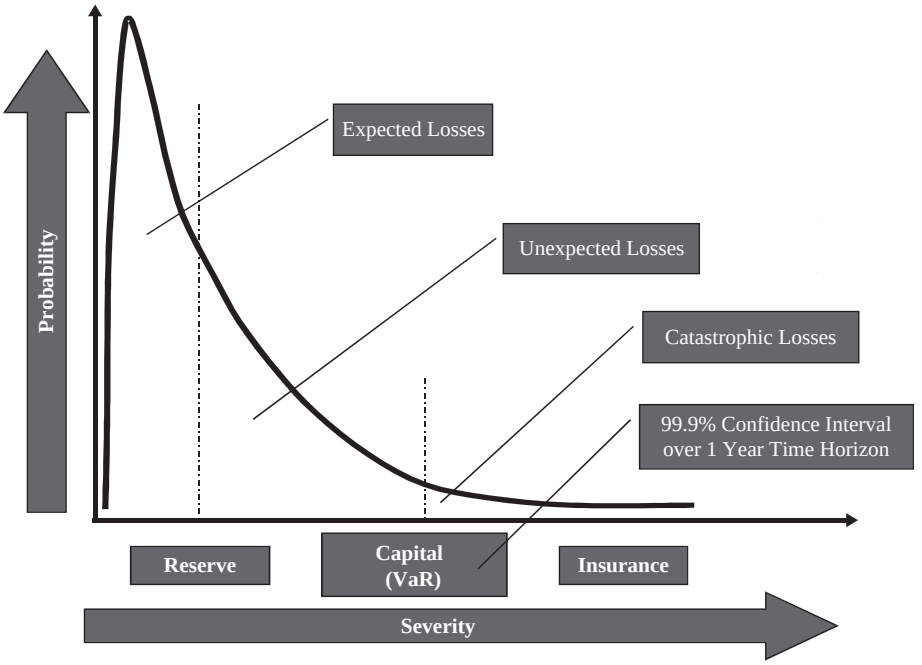


EXHIBIT 20.2 Calculating Risk Capital

- *Capital value at risk* is an estimate of the unexpected losses at a specific confidence interval over a given time frame. There are usually measures of VaR for each of the three enterprise risks—market, credit, and operational risk. Under the Basel II regime risk coverage for each type of risk is broadly defined as:
  - *Market risk*. The risk that an enterprise's tradable assets or its interest rate differential (asset-liability gap) loses value due to market price fluctuations.
  - *Credit risk*. The risk of the enterprise not receiving payment for deploying its assets.
  - *Operational risk*. The risk of loss resulting from inadequate or failed internal processes, people, and systems or from external events. (Operational risk includes legal risk, but excludes business, strategic, and reputational risk).
- *Confidence interval* is the level of risk, expressed as a confidence interval during a prescribed time period, at which the enterprise has chosen to operate. The higher the confidence level selected, the lower the probability of insolvency. For example, at a 99.97 percent confidence level the enterprise is accepting a 3 in 10000 probability of insolvency over a one-year period. Many banks using economic capital models have selected a confidence level between 99.96 and 99.98 percent, equivalent to the insolvency rate expected for an AA credit rating.
- *Insurance* is an expense item usually purchased for protecting both *catastrophic losses* and *unexpected losses*. The enterprise needs to make trade-offs between the uncertainty of capital insolvency and the costs of insurance. Insurance is usually built into the cost of the product with a commensurate effect on its profit margin.

- *Economic capital* is typically defined as the difference between some given percentile of a loss distribution and expected losses. It is the common currency for risk adjusted performance sometimes referred to as the unexpected loss measured at a specified confidence interval.

FINANCIAL ACCOUNTING VERSUS RISK ACCOUNTING

Contemporary financial and risk reporting systems are simply not equipped to provide real-time, or near real-time information on aggregated enterprise risk exposures in financial institutions. Whereas the transaction values used for financial accounting may give some indication as to credit risk exposures, that is, the size in monetary value of the credit portfolio, they provide somewhat less indication of market and liquidity risk exposures and little to no indication of operational risk exposures.

In the case of market risk, discussion is currently revolving around "mark-to-model" accounting as an incremental step to "mark-to-market" accounting as a means of bringing financial reporting closer to a disclosure of true risk exposures. There are some issues here. Risk models suffer from the relative subjectivity of inputs related to scenario analysis and stress testing and there are ongoing concerns as to the relevance, completeness, and quality of the underlying data.

Operational risk is even further behind as the industry still hasn't resolved the conundrum of how to identify a financial institution's complete portfolio of operational risk exposures and put a consistent and comparable value on them. Under such conditions, aggregating enterprise exposures within a common measurement framework across all the risk categories appears an unachievable goal. New thinking is required!

The primary source of inputs to financial accounting systems is transactions that are uniquely coded to ensure their correct accounting and reporting in financial and management accounts. Could these same transactions also be uniquely coded for risk accounting? This is the question that needs to be answered in the context of creating an integrated risk framework. The transaction values used for accounting purposes are simply not suitable for the real-time or near-real-time reporting of risk exposures. The conclusion is that a new unit of enterprise risk exposure measurement is required to complement the monetary transaction values used for financial accounting. Thus, the *cross-enterprise solution* first creates a new unit of financial risk exposure measurement, the *enterprise risk unit (ERU)*. Integrated enterprise risk solutions will evolve around this new risk currency, the ERU, in the same way that market risk evolved around a standard, VAR; commercial credit risk evolved around credit ratings; and retail lending practices evolved around credit scores.

10 PRINCIPLES OF EFFECTIVE ENTERPRISE RISK MANAGEMENT

It is important to identify the requirements of an effective integrated cross-enterprise solution that includes the monitoring and management of risks denominated in ERUs. The result is these 10 principles:

Risk Monitoring

1. A financial institution's total enterprise risk exposure is measured by applying a common measurement framework to all the transactions that comprise an operating universe.
2. A standard measurement unit of enterprise risk exposure must have a meaningful and relevant additive value that correlates with transaction values and risk drivers.
3. The management of enterprise risk exposures is most effective when it is applied at the precise moment such risk exposures are created.
4. Enterprise risk exposures are created upon transactions entering the operating environment and each time amendments or enrichments to those transactions are made.
5. An effective enterprise risk exposure monitoring system reports on the status of causal factors (key risk indicators and key operating performance indicators) and corresponding risk exposures relative to a comprehensive set of predetermined operating parameters (risk appetite).
6. Risk management occurs when business judgment is applied in response to the risk exposures reported by an effective risk monitoring system.
7. Effective mitigation of enterprise risk exposures requires the status of causal indicators to be calibrated relative to formally adopted risk management best practices and/or conditions (benchmarks).
8. A positive risk culture results when incentive and reward programs are linked to risk adjusted performance measurements derived from an effective enterprise risk exposure monitoring system.

Risk Management

9. An effective enterprise risk management system must have the ability to systematically link negative outcomes (losses) with related risk exposures which, in turn, reflect the prevailing status of all relevant causal indicators.
10. An effective risk management system is complete when enterprise risk exposure measures become predictive through their ongoing statistical correlation with actual monetary loss experience.

A TRANSACTIONAL APPROACH

Similar to financial accounting systems, the risk element in integrated risk solutions should be applied at the transaction gateway on the principle that cross-enterprise risk exposures are triggered the moment a financial institution accepts transactions into its operating environment.

Enterprise risk exposures become theoretically known to an institution as transactions are entering its operating environment. Here monetary values can be attached to them, their risk relevance with respect to each risk category (operational, credit, market, and liquidity) assigned and the status of risk mitigation systems relative to best practices observed. These constitute "internal" factors as they can be directly managed and influenced by an institution to either increase or decrease the amount of exposure to risk.

For example, a transaction's essential characteristics will determine whether it has credit risk relevance. If the institution implements a best practice "flawless" credit risk management framework, then it can be assumed that the "internal" exposure to credit risk will be fully mitigated. But if the risk management framework is flawed, that is, less than best practice, then the amount of internal exposure to risk will in some way be related to: (1) the degree of credit risk relevance of the transaction (exposure driver); (2) its monetary value (value driver); and (3) the extent to which the risk management framework is best practice (risk mitigation driver).

There are also "external" factors that create exposure to credit risk that are beyond an institution's direct management and influence. These can relate to, for example, changes in an obligor's credit rating or in macroeconomic factors affecting credit quality. Institutions typically employ statistical techniques to measure their external (unknown) exposures to risk. This is the VaR, which, for the three principal Basel II risk categories (credit, market, and operational), is calculated by reference to relevant historical data and scenario analysis at a 99.9 percent confidence level.

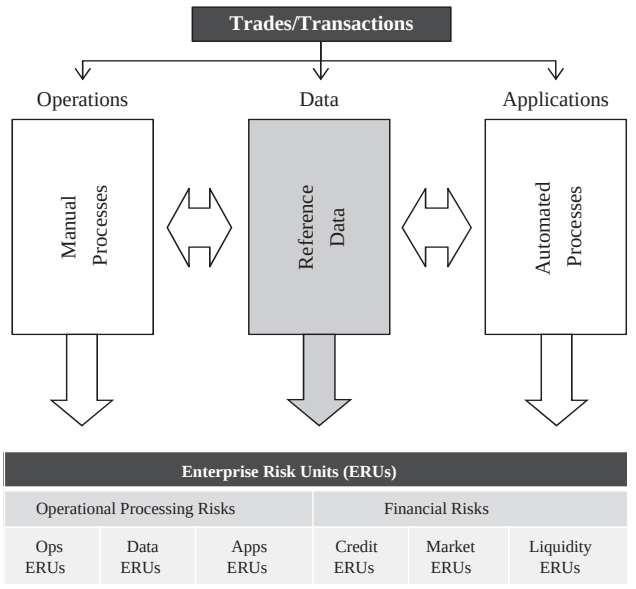
But if internal "known" exposures to risk are to be aggregated across all the risk categories there needs to be a standard unit of measurement that is additive and can be applied to enterprise risks. The essential question is whether the three enterprise risk exposure drivers referred to previously (exposure, value, and risk mitigation) can be combined within a common measurement framework to create such a standard unit of exposure measurement—the ERU?

In constructing the ERU we offer the basic premise that transactions drive enterprise risk exposure. Banks construct operating environments comprised of people, technology, facilities, processes, and controls to handle transactions and reduce cross-enterprise risk exposures which, to a point, increase as the volume and relative complexity of transaction throughput increase.

Operating environments can be deconstructed into a simple model represented by three key operational pillars—people, data, and systems. A bridge between operational metrics and cross-enterprise risk management can be constructed in the form of a common risk measurement framework. Here, we observe a "normalized" measure of cross-enterprise risk exposure, that is, the ERU, and tie it to both the operating processes and to the financials of a financial institution.

The theoretical flawless interaction of the three operational pillars (manual process, automated process, and data) produces zero risk and is represented as an operating environment with 100 percent straight-through-processing (STP) at 100 percent best practice (BP) as the benchmark. By measuring the level of failure of those interactions in ERUs we are able to "rank order" the outcomes, and assess relative risk through both a set of quality indices (%BP) and benchmarks. Thereafter, the risk measures are correlated with loss history, as was demonstrated over time in the development of credit risk measures, and the usual measures of VaR are calculated. Exhibit 20.3 depicts the transactional view of enterprise risk and the association with ERUs.

In general, operational sophistication increases as transaction volumes increase primarily due to enhanced automation. The relative quality and effectiveness of risk mitigation also increase as transaction volumes increase. The net result is that the rate at which operational risk exposure is created decelerates relative to the rate at which transaction volumes increase. Therefore, an approach to measuring operational risk recognizes this relationship and progressively reduces the rate



Enterprise Risk is the consequence of the failed and/or insecure interaction of manual processes (Operations) and automated processes (Applications) with Data relative to the processing of transactions (operational processing risks) and the management of credit, market and liquidity exposures (financial risks)

Enterprise Risk Units (ERUs) is the standard unit of measurement applied to enterprise risk exposures

EXHIBIT 20.3 Definition and Source of Enterprise Risk and Enterprise Risk Unit (ERU)

at which risk exposure is valued relative to increased transaction volume. (A further discussion of this relationship is provided in the next section and depicted in Exhibit 20.8.)

Financial transactions can be thought of as a set of computer-encoded data elements that collectively represent:

- Standard reference data, identifying it as a specific product or tradable instrument defined by its initial offering terms and conditions, and bought and sold by specific identified counterparties and/or their beneficial owners.
- Variable transaction data such as traded/purchased date, quantity, and traded/purchased price.
- Associated referential information such as credit ratings, standard payment and settlement terms and instructions, corporate action information, and so on.

The reference data components of a financial transaction identify it as a specific financial product (product/security number, symbol, market, etc.), its unique type, terms and conditions (asset class, maturity date, conversion rate, etc.), its manufacturer or supply chain participant (counterparty, dealer, institution, exchange, etc.), its delivery point (delivery, settlement instructions and location), its delivery or inventory price or balance (closing or settlement price), its market reference prices (last sale, bid/ask quote), and its currency. Analogous to specifications for manufactured products, reference data also defines the products' changing specifications (periodic or event driven corporate actions), occasional changes to sub-components (calendar data, credit rating, historical price, beta's, correlations, volatilities) and seasonal incentives or promotions (dividends, capital distributions, and interest payments).

Transactions fail if data is faulty or the data recorded in sending and receiving systems are inconsistent and can't be matched. Regulatory and compliance failures result if supply chain or product reference data do not contain the correct reporting classifications. Financial accounting and reporting processes fail if account and cost center codes are faulty or are not correctly specified in transactions and reporting matrices. Losses of revenue can occur if sales volume or particular trades are incorrectly valued due to faulty price and rate related data.

CROSS-ENTERPRISE SOLUTION

In conceptualizing a cross-enterprise solution, there are parallels with financial accounting systems. In principle, financial accounting systems start with transactions that are uniquely coded so that they can be directed to the appropriate general ledger accounts and cost centers. Various tables and templates are created and maintained by financial controllers, the most important being the standard chart of accounts, to drive the financial and management accounting and reporting processes.

Cross-enterprise risk management merely augments transactions with coding that steers them through tables and templates created and maintained by risk management to drive risk reporting processes. Consequently, a cross-enterprise risk system is a "risk accounting system" that runs alongside financial accounting systems.

Exhibits 20.4 and 20.5 are a diagrammatic representation of a risk accounting system that has been designed to produce enterprise risk exposure reporting based on the ERU. Exhibit 20.4 demonstrates how production systems are interfaced to a risk metrics server, and Exhibit 20.5 illustrates the tables and templates supported by the risk metrics server that calculate the ERUs by risk category.

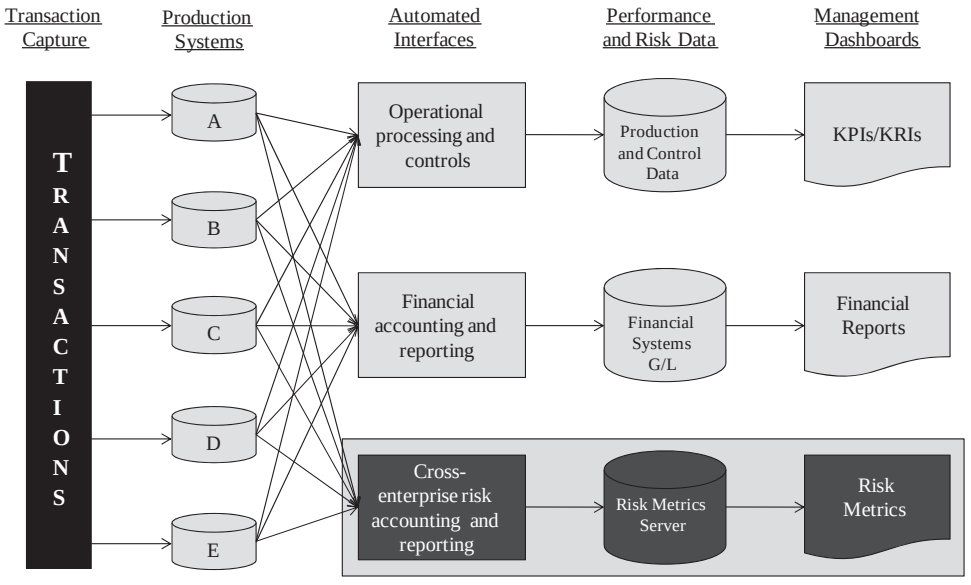


EXHIBIT 20.4 Production Systems Interface with Risk Metrics Server

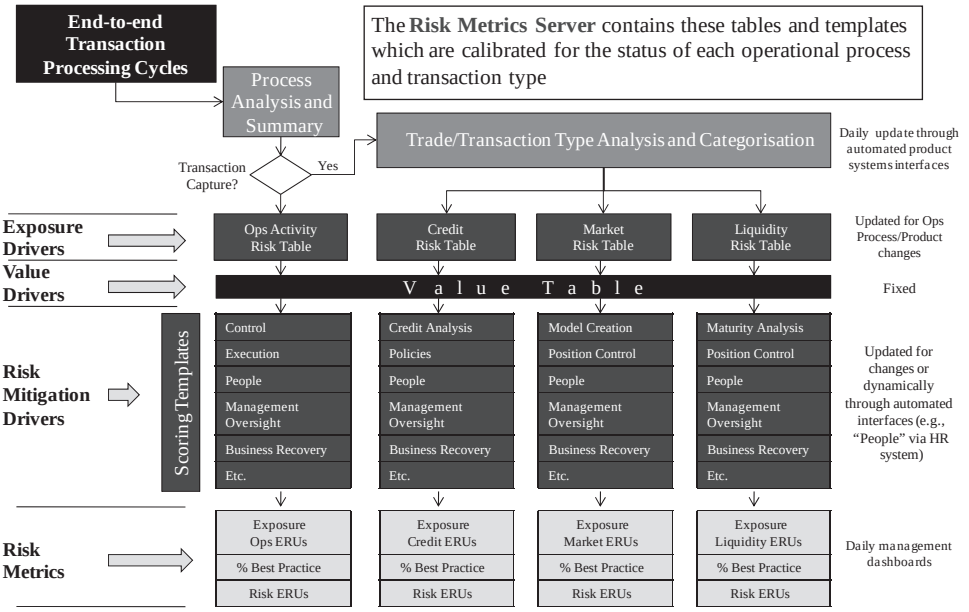


EXHIBIT 20.5 Risk Metrics Server Tables and Templates

The risk metrics server contains three types of tables/templates:

1. Risk tables—exposure drivers
2. Value table—value drivers
3. Key risk category (KRC) scoring templates—risk mitigation drivers

A fourth table (see Exhibit 20.6) is available that is dynamically updated from operational metrics (key risk indicators [KRIs] and key performance indicators [KPIs]) provided from source systems. These metrics are, in turn, translated into key risk category (KRC) scoring templates and converted into delineated relative value risk weightings and, thereafter, prorated against the fixed intervals between 0 and 100, as depicted in Exhibit 20.8.

Risk Tables (RTs)

An extract from the Ops Activity Risk Table relating to Payments and Settlements is shown in Exhibit 20.7. The RTs are comprised of preidentified process/product characteristics by risk category with a risk weighting attached. For example, if a new debt instrument has been approved for trading, the processes that comprise the end-to-end transaction processing cycle for the new product will be mapped to the Ops Activity Risk Table and the weightings accumulated according to the relative risks of the operational activities performed.

If a process involves transaction capture, amendment, or enrichment the other risk tables (credit, market, and liquidity) are triggered. If it is a traded product then the transaction will have been precoded to pass through the Market Risk Table, and

Metric Category ID	AMT/ VAL	KRC Template Group			Template Details			Formulas Using Multiple Metric Category ID
		Ops	Data	Apps	Name	#	Cell	
General Ledger (GL)	*No. a –n*	*x*	*x*		People	A001	A1	
Transaction Counts (TC)	*No. 1 –n*	*x*						
P&L Category (PL)	*x,y*	*x*						
Balance Sheet Item	*x,y*							
Cash Flow Statement	*x,y*							
Fund Sources and Uses	*x,y*							
Capital	*x,y*							
Shareholders	*fa(v)*							
Market Metrics	*fa(w)*		*x*		Quality Data	B001	A1	STP Rate = Reconciled Items / Total Records
Human Resources	*fa(x)*	*x*	*x*	*x*				
Pension	*fa(y)*							
Taxes	*fa(z)*							
Regulator/Compliance	*s,t*				Business Recovery	G001	C3	
Professional Services	*s,t*							
Availability/Usage	*s,t*			*x*				
Fixed Assets	*s,t*	*x*						

Formula Metrics

GL a+GL b
TC 1+2… +n
PL ?(x,y)+/-(xn,yn)

Metric Category ID—The originating source named file/other system ID

Legend:

No. a –n	General Ledger/Sub Ledger Account Number
No. 1 –n	Assigned Number/Code for Transaction Counts
x,y	Row, Cell of spreadsheet data
fa(v) – fa(z)	Formula for computing metric AMT/VAL

AMT/VAL—Amount in currency, transaction value count in units

Key Risk Category (KRC) Template

Group		
Ops	Template KRC Series Number 1–1	relating to Manual Operational Process
Data	Template KRC Series Number 1–m	relating to Data related Processes
Apps	Template KRC Series Number 1–n	relating to Software related\business Processes

Template Name / #—Unique Name (i.e., Quality Management, Business Recovery, et al.) and number assigned to each Scoring Template.

Cell—Location coordinate within each Scoring Template, where multiple scores are developed for each KRC Template.

Formulas Using Multiple Metric Category IDs—More granular logic to accommodate Cell-level Scoring. Template metrics calculations. Example shown of calculation of Straight-Through—Processing Rate for Quality Data Template.

Note—Designations in boxes such as "x," "People," "B001," "C3" etc., are there for illustrative purposes only.

EXHIBIT 20.6 Dynamic Scoring Table

risk weightings will be accumulated for factors such as the maturity of the product (a new product attracts a higher weighting), complexity, market liquidity, and so on. The product will attract further weightings if it has credit risk or liquidity risk relevance.

The risk tables are set for each product/transaction type and are updated by risk management whenever there are product or process changes.

EXHIBIT 20.7 Extract from Ops Activity Table—Payments and Settlements

Description	Activity Risk Weighting
Release value items (including standard settlement instruction and standing order/direct debit maintenance) to guaranteed counterparties	2
■ Intercompany and intracompany ■ Guaranteed settlement (e.g., central exchanges/Continuous Link Settlement) ■ Delivery versus payment agreements	
Release value items (including standard settlement instruction and standing order/direct debit maintenance) to financial market counterparties	5
■ Banks and other financial institutions	
Release value items (including standard settlement instruction and standing order/direct debit maintenance) to other parties	10
■ Nonfinancial market counterparties ■ Third parties	

Value Table (VT)

The VT is shown in Exhibit 20.8 and is a logarithmic curve that depicts the relationship between transaction values and risk, that is, the marginal increase in risk reduces as transaction (processing) values increase. Transactions are categorized and grouped on a daily basis and are mapped to the value table and the applicable value band weighting is extracted. Depending upon the granularity desired the VT can be recalibrated to fit smaller-size organizations and can be related not only to revenue but to position value.

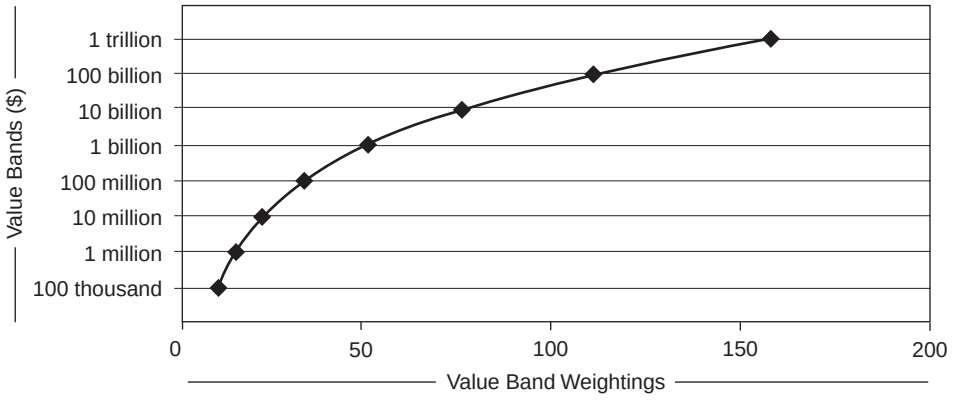


EXHIBIT 20.8 Value Table

Key Risk Category (KRC) Scoring Templates

Two sample KRC scoring templates are shown in Exhibit 20.9 relating to *execution* (benchmark based) and *business recovery* (best practice statement based). Each template is scored whereby each score represents the actual status relative to best practices. Scores are updated upon changes or dynamically through automated

Ops Key Risk Categories/Weightings		
	Weighting	Score
Control	10	0 to 100
People	10	0 to 100
Execution	10	0 to 100
Business Recovery	8	0 to 100
Risk Management	6	0 to 100
Management Oversight	6	0 to 100
Application Security	4	0 to 100
Physical Security	4	0 to 100
Policies and Procedures	2	0 to 100

Execution: levels of automation vs. manual workarounds; levels of repair rates; and the stability of core application(s).

Level of automation or STP rate:
- 100% score 100 (Best Practice)
- 75% score 75
- 50% score 50
- 25% score 25
- 0% score zero

Average percentage of input rejection/repair:
- 0% score 100 (Best Practice)
- 5% score 75
- 10% score 50
- 25% score 25
- 50% score zero

Number of core system failures in year:
- None score 100 (Best Practice)
- 1 score 75
- 2 score 50
- 4 score 25
- > 12 score zero

Business Recovery: continuation of operations at an alternative site in a time frame that is acceptable

Best Practice score 100

Deduct following scores from Best Practice score if statement does not apply:
- Recovery or reactivation at alternative site in acceptable time frame (100)
- Formal business recovery plan (100)
- End-to-end disaster simulation (75)
- Plan complete and comprehensive (30)
- Supervisory review of plan (20)
- Key employees fully briefed (15)
- Key employees active participation in disaster simulation (10)
- Business recovery specialist review of plan (10)
- Key employees' contact details current (5)
- Notification test performed (5)
- Key employees ready access to offsite copy of plan (5)

EXHIBIT 20.9 Key Risk Category Scoring Templates "Execution" and "Business Recovery"

Key Risk Categories	Control Evaluation	People	Execution	Business Recovery	Risk Culture/ Management	Management Oversight	Application Security	Physical Access	Policies & Procedures	% Best Practice	Risk	Exposure
Category Weightings	10	10	10	8	6	6	4	4	2		ERUs (Thousands)	
Transaction Category A												
Type 1	25	50	45	15	50	75	75	100	50	**47.8**	**86**	**165**
Type 2	80	100	50	0	30	50	40	100	20	**56.3**	**48**	**110**
Type 3	25	50	45	15	50	75	75	100	50	**47.8**	**57**	**110**
Type 4	0	30	25	5	40	10	70	100	0	**26.2**	**111**	**150**
Trans Cat A—%BP	29.3	54.7	40.4	9.1	43.1	51.6	66.4	100.0	29.8	**43.5**	**302**	**535**
Transaction Category B												
Type 1	70	70	50	100	100	100	70	100	100	**79.7**	**18**	**90**
Type 2	70	70	50	100	100	100	70	100	100	**79.7**	**20**	**100**
Type 3	70	60	85	80	60	85	75	100	80	**75.3**	**54**	**220**
Trans Cat B—%BP	70.0	64.6	68.8	89.3	78.5	92.0	72.7	100.0	89.3	**77.3**	**93**	**410**
Total—%BP	47.0	59.0	52.7	43.9	58.5	69.1	69.1	100.0	55.6	**58.2**	**395**	**945**

EXHIBIT 20.10 Sample Scorecard

interfaces (e.g., people scores via the human resources system). KRC scores are blended with other weightings:

1. KRC weightings which are calibrated according to the relative risk mitigation impact of each KRC.
2. The ERUs representing risk weighted transactions that interact with the KRC. From these inputs we can calculate the risk metrics using the formulae below where W = weightings and S = scores.

- Exposure ERUs (ExpERU) = $RT^W \times VT^W$
- % Best Practices (%BP) = $\frac{\sum (KRC^S \times KRC^W \times ExpERU) \times 100}{\sum (100 \times KRC^W \times ExpERU)}$
- Risk ERUs (RiskERU) = $\frac{(100 - \%BP) \times ExpERU}{100}$

A sample scorecard demonstrating the calculation of Operations ERUs and %BPs has been reproduced in Exhibit 20.10.

The cross-enterprise process described represents a risk accounting approach similar to financial accounting systems, as it is transaction based. Transactions are captured, categorized, translated into a common currency "ERU" and posted to "risk accounts" by passing them through tables and templates owned and maintained by risk management. In this way, risk metrics can be consolidated and aggregated for reporting via management dashboards by transaction category, organization, geography, risk type, key risk category, and so on. This process also incorporates a

budget module so that risk appetite can be denominated, allocated, and monitored in ERU and %BP.

PREDICTIVE RISK MODELS

Exhibit 20.11 illustrates the functioning of the cross-enterprise predictive risk model. Insofar as the cross-enterprise process is transaction based, risk metrics (ERUs and %BPs) are permanently attached to each transaction upon entry into the operating environment. When loss events occur they are in turn mapped to the transaction(s) or groups of transactions that relate to each loss event. In this way, monetary losses are linked to risk exposures (in ERUs) and causal factors (in %BPs) which facilitates statistical correlation. As loss data is gathered and correlated with related exposure and causal data, daily transaction based ERUs will become increasingly loss predictive.

In the Cross-Enterprise risk modeling process, internal loss data that has been enhanced by the attachment of context information (ERUs and %BPs) are further complemented by external loss data that can be accessed through consortia such as ORX.[12] The application of stress testing and scenario analysis completes the risk modeling process.

The Cross-Enterprise process is superior to other solutions currently available as it is the only solution that includes real-time or near real-time transaction-based exposure and risk information in its predictive modeling. This is discussed in more detail in the following section.

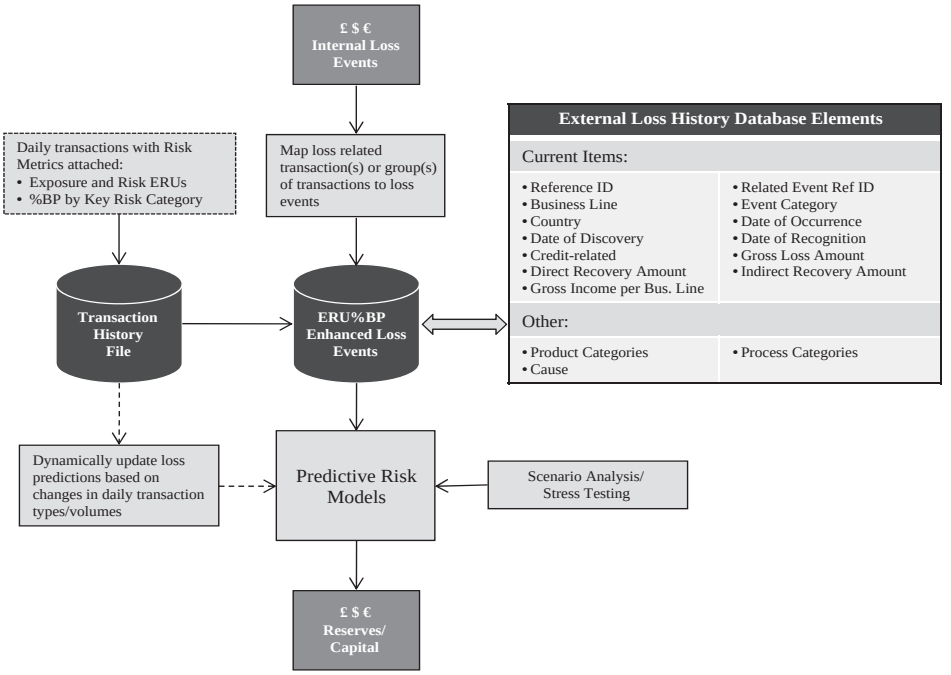


EXHIBIT 20.11 Cross-Enterprise Predictive Models

CONVENTIONAL SOLUTIONS VERSUS CROSS-ENTERPRISE PROCESS

Conventional operational risk management solutions are generally component based and are typically comprised of:

- An internal loss event data collection tool that enables the collection, classification, and maintenance of operational risk loss events.
- Action plans that can be created for loss events with specific workflows for the acceptance of loss data and the execution of action plans.
- A risk and control self-assessment (RCSA) tool that allows firms to inventory key risks and controls and then make decisions to control/mitigate risks.
- A key risk indicator (KRI) tool that enables the identification of key risks and associated risk thresholds so that a firm can monitor values and identify trends that might lead to unacceptable risk.
- A scenario analysis tool that is designed to identify, arrange, and present information including internal loss data, relevant external loss event information, and key risks and controls identified during the RCSA for scenario analysis.
- A capital modeling tool that provides data analysis capability combined with sophisticated tools for modeling loss events.

If exposure to risk exists, it follows that the occurrence of loss events is inevitable. In these circumstances if managers are to exercise their business judgment effectively they must have access to enterprise level real-time or near-real-time information on the size and distribution of risk exposures in a form that they can analyze and drill to the causes.

Exhibit 20.12 provides a comparison of conventional methods and the cross-enterprise process applied to operational risk. Because the cross-enterprise process uses a standardized additive unit of risk measurement (the ERU), which is comparable within and between financial institutions, benchmarking reports and management dashboards are available that conventional solutions are unable to produce. Examples of benchmarking reports and management dashboards are provided in Exhibits 20.13 and 20.14.

Conventional solutions rely almost exclusively on qualitative risk management mechanisms such as KRIs and RCSAs. Such devices are unquestionably valuable but suffer from their inherent subjectivity. Even in the case of KRIs, line managers generally set their own trigger or threshold points to determine the relative severity of a potential risk condition (red, amber, or green), thereby influencing how risk exposures are reported.

But the real limitation of indicators and self-assessments is that they are nonadditive and, consequently, cannot be consolidated and aggregated to provide consistent and comparable "top-down" profiles of operational risk exposure at all levels of the enterprise. This constitutes a serious impediment to the effective management of enterprise risks in financial institutions.

The absence of additive measurements of exposure to risk also inhibits the ability to apply statistical techniques to predict future losses. Statistical correlation shows whether, and how strongly, pairs of variables are related. Important in risk

EXHIBIT 20.12 Comparison of Conventional Solutions vs. Cross-Enterprise Solution

Operational Risk	
Conventional Method	**Cross-Enterprise Solution**
Toolkit: ■ Internal loss event data collection ■ Risk Control Self Assessments (RCSA) ■ Key Risk Indicator (KRI) monitoring ■ Operational capital modelling	Toolkit: ■ Internal loss event data collection ■ Value Table/Ops Activity Table/Key Risk Category (KRC) best practice scoring templates ■ Operational capital modelling
Discuss past losses (if available) with business line management	■ Chart Processes, identify and evaluate controls ■ Map Processes to Value Table and Ops Activity Risk Table ■ Determine actual and target KRC scores
Report current losses to risk manager who records losses in loss event database	Record losses upon occurrence
Ask operating management to estimate frequency and severity of future losses at 99.9% confidence level	Calculate actual and target ERUs and % Best Practice (%BP) for each Process and agree/log close-the-gap actions
Present loss, frequency estimates vs. historical losses report	Present "ERU & %BP Risk Reports" and "Route Map to Operational Excellence"
Discuss projects for risk mitigation	Monitor status of actions
Agree on project/cost/time frame and loss history event removal for capital reduction benefit, review progress and if project complete remove/reduce loss event	Confirm completed actions and recalculate ERUs and %BP
Roll-up estimates of frequency and severity by business line	Roll-up ERUs by Business Line and compare to total firm losses at 99.9% confidence
	Report ERUs to External Loss Databases, trade associations' benchmarking services
Incorporate external loss data into scenario analysis, stress testing	Incorporate external loss data and ERU benchmarks into scenario analysis and stress testing
Calculate OpVaR by Business Line via analysis at the 99.9% confidence level	Calculate Ops Value-at-Risk (OpVaR) by Business Line via proportionality to total firm's ERUs and analyse at 99.9% confidence level
Discuss with operating management to determine cause of loss event and preventive measures	Drill into Dashboard, review external benchmarking data and present to operating management

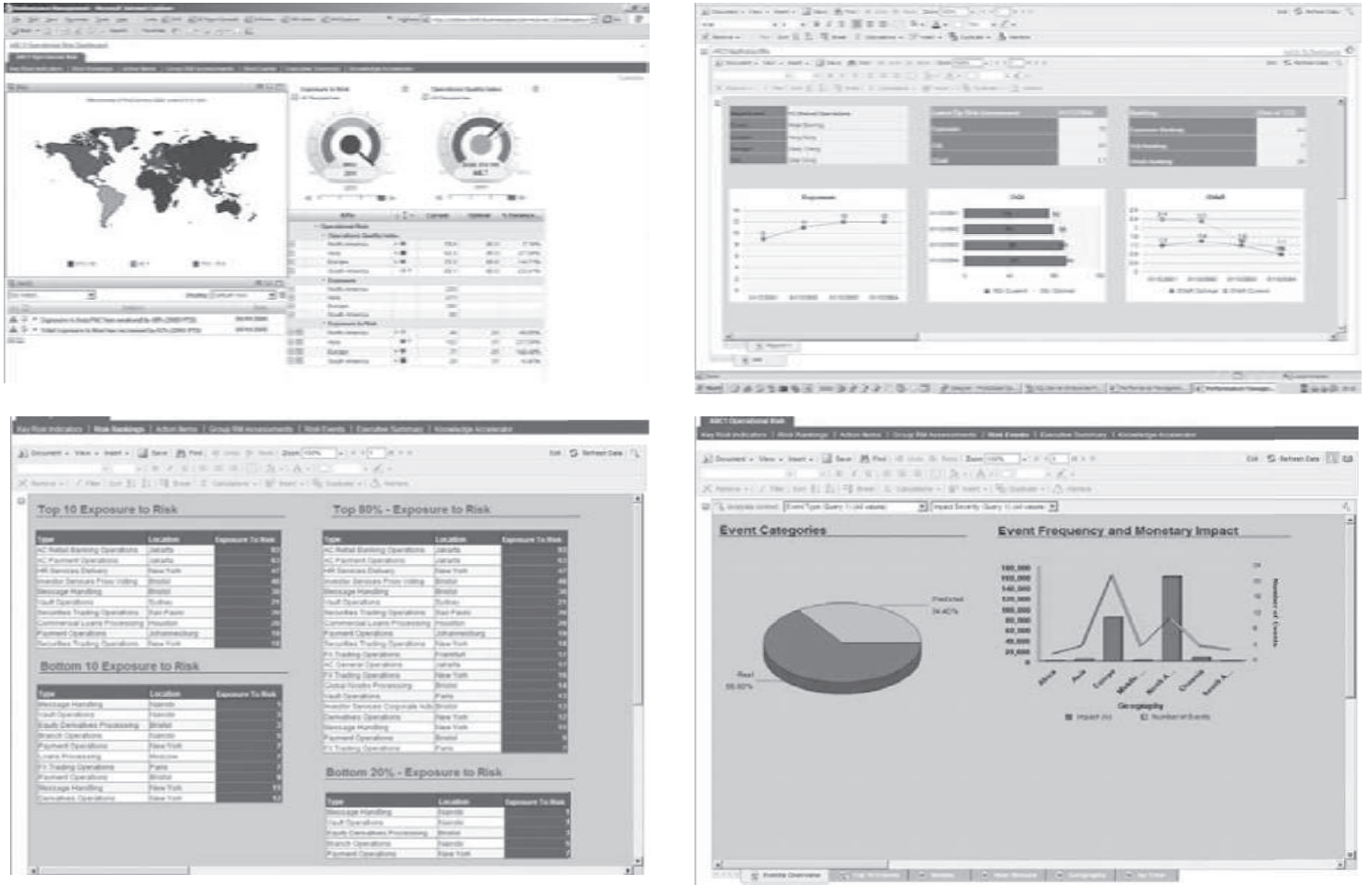


EXHIBIT 20.13 Sample Enterprise-Level Operational Risk Dashboards Based on %BPs and ERUs
Source: Courtesy SAP.

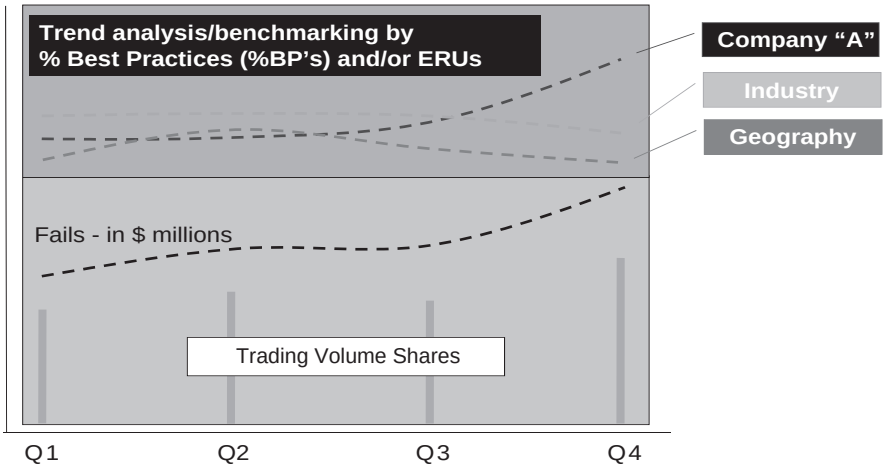


EXHIBIT 20.14 Sample Industry Level Benchmarking Report

management is the correlation between exposure (the total risk-weighted size of transactions) and risk (the probability that risk mitigation is ineffective causing a possible negative outcome, i.e., losses). By analyzing how actual losses occur relative to measurements and distributions of risk and exposure, risk managers can fine-tune their risk models and continuously update their predictions of future losses consistent with changes in risk and exposure.

Statistical correlation requires the ability to perform in-depth analysis of the size and distribution of enterprise risk exposures and a store of historical loss data with each loss event linked to the status of exposure and causal factors at the time the loss event occurred. The permanent attachment of such context information to each loss event is important in enterprise risk management where context, and consequently exposure to risk, is constantly changing due to, for example, fluctuations in transaction volumes, organizational changes, new technology and systems, new products, new regulations, and changes in operating processes.

Because the outputs of KRIs and RCSAs are, by their very nature, subjective and not expressed in value-bearing units of measure, their correlation with actual loss experience is severely inhibited.

CONCLUSION

The rapid adjustments that regulators are now making to the basic elements of the Basel II accord should further reinforce the inextricable movement away from high-flying, take-the-money-and-run mind-sets to a more responsible intertwined global financial services industry. Here, the dominant paradigm shift will be found in embedding a risk culture into the very nature of performance metrics and incentive compensation schemes. Without such a shift there can be no true management of risk as being a discipline that leads proactively down a path of risk mitigation. Instead, it will continue to be a retrospective-leaning discipline, relying on loss history as a

basis to predict future loss experience. In the absence of true risk management, the short-term exercises in risk assessments and capital quantification will continue with periodic blowouts, as has been the norm in this 20-year period of experimenting with the discipline of risk management.

The transaction-based cross-enterprise risk system described here is more than a risk management system; it truly is a risk-adjusted performance measurement system. It may one day take its rightful place alongside management information systems, which had a two-generation gestation period before management was able to see reliable customer, product, and business unit performance and cost attribution data.

Today, we are at the early stage in taking the next leap forward in management information systems—tagging each transaction with its associated "riskiness." We similarly tagged a customer, product, or business unit with its associated profit or loss data to drive the financial and management accounting and reporting of the company.

Categorizing the riskiness of transactions that enter the processing streams of financial institutions and aggregating them throughout the many silos that now characterize the organizational structures of many of them to ultimately find their way into capital calculations will be a daunting task. Indeed, it will not be dissimilar to the two-generation period of evolution that passed before we got management information systems right. In risk management we are already past the first generation mark. We believe the next generation will bear breakthroughs in theory and in the practical joining of operational metrics with risk metrics. In this chapter, we have humbly offered our views as to how this may be achieved.

NOTES

1. Basel Committee on Banking Supervision, "Operational Risk" (consultive paper)—January 2002, p. 2.
2. Citigroup, "Citi Reports Fourth Quarter Net Loss of $9.83 Billion, Loss per Share of $1.99" (press release), January 15, 2008, p. 12.
3. Claudio Borio, Monetary and Economic Department, Bank for International Settlements, BIS working papers No. 251, "The Financial Turmoil of 2007: A Preliminary Assessment and Some Policy Considerations," March 2008, p. 28.
4. Société Générale, General Inspection Department, Mission Green Summary Report, May 20, 2008, p. 2.
5. Basel Committee on Banking Supervision, "Cross-Sectoral Review of Group-wide Identification and Management of Risk Concentrations," April 2008.
6. Bank for International Settlements (BIS), Conference Paper, November 1999; www.bis.org/list/bispapers/from_01011998/index.htm.
7. Federal Reserve, May 24, 2007, "Response by the Advanced Measurement Approach Group of the Risk Management Association to the Proposed Supervisory Guidance for Internal Ratings-Based Systems for Credit Risk, Advanced Measurement Approaches for Operational Risk, and the Supervisory Review Process (Pillar 2) Related to Basel II Implementation," p. 8.
8. "Proposed Supervisory Guidance for Internal Ratings-Based Systems for Credit Risk, Advanced Measurement Approaches for Operational Risk, and the Supervisory Review Process (Pillar 2) Related to Basel II Implementation," *Federal Register* 72(39) (February 28, 2007): 9170.

9. Ibid., p. 9173, footnote 13.
10. Robert L. Burns, "Economic Capital and the Assessment of Capital Adequacy." *RMA Journal* (April 2005).
11. Risk-Based Capital Guidelines; Capital Adequacy Guidelines, Standardized Framework; Proposed Rule and Notice; Office of the Comptroller of the Currency, Treasury; Board of Governors of the Federal Reserve System; Federal Deposit Insurance Corporation; and Office of Thrift Supervision, Treasury, Joint Notice of Proposed Rulemaking, *Federal Register* (June 26, 2008).
12. ORX, The Operational Riskdata eXchange Association (www.orx.org).

CHAPTER 21

Throughput Accounting

Chris Zephro

BACKGROUND

One of the key responsibilities of a finance department is to ensure that managers are making profitable decisions for the company that balance risk and opportunities. Most managers are given templates to fill out to evaluate how a decision will impact things like return on investment (ROI), net present value (NPV), payback period, and other commonly used financial measures. However, these measurements tend to be difficult to apply to everyday decisions that managers have to make. Additionally, these measurements do a poor job at looking at the holistic impact a particular decision has across the enterprise. In other words, how do we know that a decision is not improving one area of the business at the expense of the whole company and achievement of its goal?

Dr. Eliyahu M. Goldratt's groundbreaking book, *The Goal*, challenged many of the traditional assumptions of today's modern business world.[1] Goldratt proposed that a company's performance is dictated by a few, typically one, key constraint within an organization, and by maximizing the performance of the constraint, you maximize the performance of the entire organization. Goldratt called his methodology *Theory of Constraints* (TOC), and many companies have leveraged TOC to reap tremendous benefits within operations, project management, financial decision making, and risk management.

Suppose you have a basic process or manufacturing flow like that shown in Exhibit 21.1.

The process starts with Operation A, which has the capability to produce 15 units per day. Operation A then passes its processed units to Operation B, which has the capability to process 5 units per day. Finally, Operation B passes its units to Operation C, which has a capacity of 10 units per day, and then we have a finished product or service. The question is: how many units can this system produce? The answer is obviously 5 units per day, because Operation B, the constraint, dictates the entire output of the system. TOC recognizes this fact and applies an approach and metrics to ensure that the system is maximizing its output and not producing waste.

However, common measurements used by companies today, inspired by cost accounting, actually drive much inefficiency throughout the operation. For example, measurements such as *capacity utilization* and *cost per unit* would encourage Operation A to produce at its maximum capacity of 15 units, to drive down costs per unit,

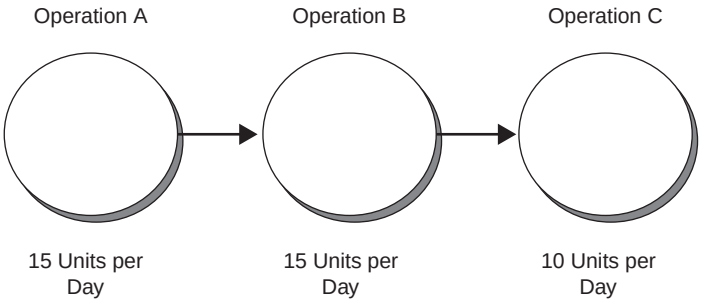


EXHIBIT 21.1 Typical Process or Manufacturing Flow

maximize absorption and improve efficiency. However, these measurements would produce excess raw material in front of Operation B and would actually increase the real costs, responsiveness, and profitability of the company. Additionally, Operation C would be penalized based on its inability to meet its metrics due to the fact that it is upstream of the constraint. Throughput accounting bridges this gap.

THE FIVE FOCUSING STEPS

It's difficult to discuss throughput accounting without providing background on what it's based on, the Five Focusing Steps of Constraint Management.[2] In order to ensure a "process of ongoing improvement" and to focus companies on the proper leverage points that govern improvement of a company holistically, Goldratt created the Five Focusing Steps:

1. *Identify* the system's constraint.
2. Decide how to *exploit* the system's constraint.
3. *Subordinate* everything else to the above decision(s).
4. *Elevate* the system's constraint.
5. If in the previous step, a constraint has been broken, *go back* to step 1.

The first step requires a company to *identify* its system's constraint. This can be done by asking the following questions:

- What limits the system performance now?
- Is the constraint inside the system (a resource or a policy) or is it outside the system (the market, material supply, a vendor, etc.)?

An easy way to identify a system constraint is to ask an expeditor where they always have to go. In a manufacturing setting, a system constraint can usually be found by walking the production floor; the constraint will have a large pile of work-in-process in front of it. However, the use of basic mathematics is always the best way to show where a constraint lies. A constraint is any resource or process that has more demand than capacity.

The next step, *exploit,* means to get the most out of the constraining element without additional investment. The goal here is to change the way you operate so that maximum financial benefit is achieved from the constraining element. One way

to achieve this is to understand the sales mix that maximizes the constraint and shape demand towards those products or services. Another way to exploit the constraint is to make sure that it is always working and not sitting idle waiting for material or operators to arrive.

Next the company must *subordinate* everything else to the decisions made in the *exploit* step. This includes making sure that parts of the system that are *not* constrained do whatever they can to support the decisions made in the *exploit* step. In other words, all nonconstraints recognize that their own efficiency is not as important as supporting the system constraint. In Exhibit 21.1 this would mean that Operation A does *not* produce to its maximum capacity of 15 units per day. Instead, it would produce something like 7 units per day (appropriate buffers would have to be calculated), which equals the 5-unit capacity of Operation B plus a buffer of 2 units to ensure that Operation B is not starved of work.

The next step, *elevate,* is required if the ROI gained by increasing the capacity of a constraint is high enough to justify the investment. Throughput accounting defines ROI as the change in net income as a result of a given investment. It is very important that companies first predict where the future constraint will be after they elevate and its impact on the global performance. Companies must also ask themselves where the constraint will go next and how difficult will it be to manage it if it shifts to a new location in the process.

It's important to understand that identification of a company's system constraint is in fact a strategic decision. I like to tell people that constraints are not bad, they just are, so you can either manage your constraint or your constraint will manage you. If your constraint continually shifts, that is indicative of a company that is not following its *exploit* and *subordination* plan and will result in a company behaving like a dog that continually chases its tail around in a circle. However, if you put the strategic decision where you want your constraint to be and properly *subordinate* and *elevate* to ensure that the constraint doesn't move to an area in your operation where you don't want to it, then you can focus your design and capacity planning around your constraint.

The final step, *go back to step 1* is the basis for Goldratt's calling TOC a *process of ongoing improvement*. This step ensures that companies move forward, rather than make a single improvement only to go back to the old way of doing business.

THROUGHPUT ACCOUNTING

Throughput accounting is a decision support tool designed to ensure that decisions, which have a financial impact take a holistic perspective and do not optimize one metric or area of the business at the expense of the whole. The fundamentals of throughput accounting were established in Goldratt's book *The Goal*,[3] and later expanded on through articles, white papers, and books dedicated to the topic.[4]

Throughput accounting differs from cost accounting in that it:

- Directly considers of the role of constraints in the financial analysis. In other words, if a decision has a positive impact at a constrained operation, throughput accounting will properly value the improvement in financial terms because it acknowledges that the constraint determines capacity, hence profit potential of the company.

- Determines profitability at the system level, instead of gross margin analysis at the product level.
- Considers the production process to be a single system that must be optimized versus optimization of every component of the system.
- Assumes most production costs do not vary directly with incremental production of a single product.
- Assumes most production costs are required to maintain a system of production, irrespective of the number of units created.
- Challenges the assumption made in product costing that producing one less drive results in a proportionate drop in the amount of overhead.

ELEMENTS OF THROUGHPUT ACCOUNTING

When evaluating a decision that has financial implications, you must quantify the impact to three key measurements:

1. Throughput
2. Investment
3. Operating expense

Throughput

Throughput (T) is the rate at which an organization generates goal units. The term *goal unit* is used because not all organizations define their goal in dollars; for example, a nonprofit organization or a hospital would not have profit as its goal. However, for the purpose of this chapter, I will assume a for-profit organization, whose goal is to make money now and in the future, hence, *goal unit* is defined as incremental cash flow through sales. Throughput represents money coming into and retained by the system.

Throughput is calculated by the following equation:

$$\text{Throughput} = \text{Sales} - \text{Truly variable cost} \tag{21.1}$$

One of the elements that make throughput accounting unique is concept of truly variable cost (TVC). TVCs are those costs that vary directly and proportionally with sales volume, in other words, the costs that a company incurs to make one more product and get it to the customer. For most companies, TVC is just raw materials.

Throughput can be measured and assessed at the unit, product family, and company level in the following ways:

- *Company level.* Sales revenue minus variable cost of all sales within *all* product lines or families.
- *Product level.* Sales revenue minus variable cost of all sales within *one* product line or families.
- *Unit level.* Unit selling price minus unit variable cost of a *single unit*.

Exhibit 21.2 shows the calculations for an electronics company that sells digital cameras with the following throughput:

EXHIBIT 21.2 Digital Camera Throughput

Product	Sales Price	TVC	Throughput
Digital camera 6 megapixels	$96	$25	$71
Digital camera 8 megapixels	$176	$50	$126
Digital camera 16 megapixels	$300	$100	$200

It's important to note that throughput is recognized only once the product is received and paid for by the company's end customer.

Investment

Investment (I) is all the money spent on asset and materials used to produce those things a company intends to sell. This includes all the assets of the company, including capital (plant, property, and equipment), as well as finished goods inventory, receivables, and intangible rights. It is important to note that throughput accounting does not recognize the concept of value added. Goldratt has stated that the only time value is added for the company is when a product is sold; hence the concept of value added is an accounting illusion.

Operating Expense

Operating expense (OE) is all the money that the company spends to turn investment into throughput. Costs in this category include fixed expenses such as salaries, rent, depreciation, supplies, interest payments, carrying costs, and overhead. Operating expense is the expense that a company pays to maintain its current level of capacity. Another way of looking at OE is what costs remain after all TVC and investment are accounted for.

A very common question asked by cost accounts is: "Why is labor considered an operating expense versus a truly variable cost?" The answer is based on how employees are paid, which is a function of time, hence their time is being burdened within the product cost. This practice causes a number of distortions, which will be elaborated on later in this chapter, causing incorrect decision making. The only time that employee wages are considered a TVC is when companies pay employees on a piece/part basis. Since this practice is rarely done today, throughput accounting considers wages to be an OE.

Another reason why throughput accounting doesn't consider labor as variable cost is that making or cutting an additional product from the production schedule rarely impacts the number of employees required. Also, you must ask yourself, how many times can you send an employee back and forth to work before they get fed up and quit?

EVALUATING FINANCIAL DECISIONS

Anytime a decision needs to be made that will have a financial impact, it's imperative that decision makers take a holistic perspective by looking at the impact to T, I, and OE.

For example, suppose a company wanted to save on transportation cost by moving from air freight to ocean freight, which would result in a $1-per-unit savings in transportation cost (TVC or OE, depending on whether transportation is charged on a per-unit or container basis). On the surface, most companies would be very tempted to go with this option without quantifying the real impact to inventory, service levels, and carrying cost. Suppose this company manufactured its products in Asia and the majority of their customers were in the United States. By moving from air to ocean, the company's transportation lead-time would go from roughly five days to upwards of 32 days, which includes clearing customs. Throughput accounting forces the logistics department to quantify, in dollars, how much additional inventory would be required to support the increase in replenishment time. Additionally, the analysis will quantify the increase in carrying cost (OE) associated with the additional inventory. The analysis would go on to compare the additional dollars in I and OE versus the savings in transportation cost, thereby taking a holistic perspective to the analysis.

ROLE OF A CONSTRAINT

Let's take another example using T, I, and OE, but this time the decision will have an impact to the company's system constraint. Remember, throughput accounting is based on the *Five Focusing Steps of Constraint Management*, which starts with identification of the system constraint, so any decisions that impact the constraint positive or negative must be accounted for properly.

Let's go back to the company with the production line from Exhibit 21.1. The company has identified its system constraint, Operation B. Now suppose an engineer has found a new part that it can use for Product A that will decrease time on Operation B, thereby increasing the capacity at Operation B from five units per day to eight units per day, which will increase the system output to eight units per day. However, the new part that the engineer wants to use will increase the bill of materials (BOM), hence TVC, by $1. Should the company go with the new part?

First let's look at some more details of Product A in Exhibit 21.3.

Total demand for product A is 10, but the company can still make only 8 per day with the new part. Now let's look at the impact to T, I, and OE in Exhibit 21.4.

The result of the analysis would suggest that the company should move forward with the engineer's proposal.[5]

It's worth noting that if the engineer had used traditional cost accounting for this analysis, the proposal most likely would have been rejected. The reason is that cost accounting does not take into consideration the role of a constraint. Instead,

EXHIBIT 21.3 Sample Product A Throughput

	Price	TVC	Throughput	Number of Units Produced	Total Throughput
Product A	$50.00	$24.00	$26.00	5	$130.00
Product A w/new part	$50.00	$25.00	$25.00	8	$200.00

EXHIBIT 21.4 Sample Product A Cost

	Product A	Product A w/new part
Throughput	$130.00	$200.00
Investment	0	0
Operating Expense	0	0

it assumes that all work stations/activity drivers are equal; hence, the increase in BOM cost would have rejected the decision, since most companies prioritize cost over throughput.

The best cost accounting would have gotten the engineer was the saved hours at Operation B resulting from the new part, which would have been counted toward capital avoidance. However, in my experience, the capital avoidance number is usually never large enough to go with a decision that would increase BOM cost, so the correct decision is never made.

Remember, the goal of a for-profit company is to *make money now and in the future*, which is an assumption that throughput accounting drives. In throughput accounting, any decision that increases throughput while simultaneously decreasing investment and operating expense will get priority over a decision that just decreases investment or operating expense. That is not to say that decreasing investment and/or operating expense alone without an impact to throughput is a bad thing; as a matter of fact, it's a good thing, assuming the decision does not have a negative impact to the constraint.

APPLYING T, I, AND OE TO TRADITIONAL BUSINESS MEASURES

Assuming a for-profit organization, where both throughput and operating expense are both money, companies can use T, I, and OE and apply them to traditional business measurements.

- Net profit = T − OE
- Return on investment = $(\Delta T - \Delta OE)/\Delta I$
- Productivity = T/OE
- Investment turns = T/I

Remember that throughput can be measured and assessed at multiple levels, and therefore, net profit can be as well. You can generate a total net profit statement for the company, which will equal the same total as a company's reported income statement, but one can also look at the expected net income over the life of a program or product line.

Companies must not loss site of net profit. There is a always a fear by some within companies implementing throughput accounting that because of the use of TVC, salespeople will sell products for just over the TVC, thereby undercutting competitors that base their prices on absorption costing. This is where the net profit

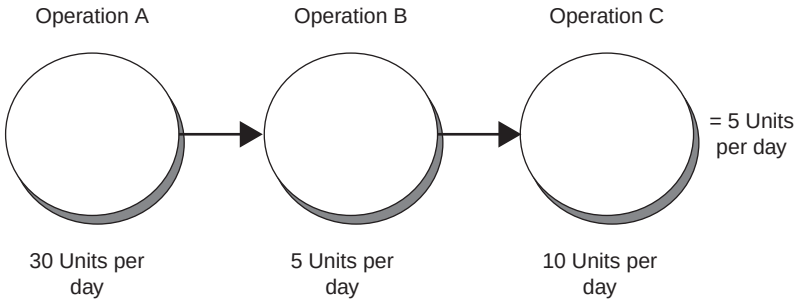


EXHIBIT 21.5 Example of Constraint on Output

calculation comes into play, because if a company did sell its products for just above TVC, they would more than likely be unable to cover their OE; hence, they would have negative net profits. It is the responsibility of sales managers to ensure that the company is cash flow and net profit positive.

ROI in throughput accounting focuses specifically on how much profit will result from a given investment not how much money will be saved by a given investment. This practice avoids making investments to nonconstrained resources. For example, it is very easy to show with cost accounting metrics such as net present value, cost per unit, payback period, and internal rate of return that buying a new Operation A machine that will increase capacity from 15 to 30 units per day is a good idea, when in fact the company is still only capable of making 5 units per day (see Exhibit 21.5).

Investment turns can be particularly useful for companies that have multiple business units with inventory specific to that business unit, as a way to understand how profit is made for a given inventory level. Suppose a company has three divisions, Retail, Distribution, and OEM, as shown in Exhibit 21.6.

Looking at Exhibit 21.6, OEM clearly has the highest throughput of all three business units; however, it achieves this with a much larger inventory level (I). This could be an indicator of a problem that would require additional analysis leveraging the TOC thinking process.[6]

PRODUCT COST—THROUGHPUT ACCOUNTING VERSUS COST ACCOUNTING

Throughput accounting values the cost of a product differently from standard cost accounting, which uses the concept of average unit cost (AUC). In standard cost

EXHIBIT 21.6 Sample of Investment Turns

	Throughput	Investment	Investment Turns
Retail	$5,000,000	$2,000,000	2.5
Distribution	$8,000,000	$3,500,000	2.2
OEM	$11,000,000	$7,000,000	1.5

accounting, AUC is calculated by a more expanded version of the following equation:

$$\text{AUC} = \text{Raw materials} + \text{Direct labor allocation} + \text{Overhead allocation} \quad (21.2)$$

Each of the allocation calculations will go up or down, depending on volume. Allocation is typically a function of the following formula:

$$\text{Allocation burden per unit} = \text{Fixed cost} + \text{Variable cost/Volume} \quad (21.3)$$

There are a number of problems that arise from this method of calculating AUC. The primary problem is volume. Based on a given volume level, the cost of the product will go up or down. This phenomenon can cause managers to build out product, regardless of actual demand, in order to spread fixed cost across a larger pool of products, thereby driving down cost per unit. This strategy, encouraged by individual factory utilization and cost-per-unit metrics, can cost companies millions of dollars in excess inventory, both finished goods and work-in-process, if their build-ahead does not match demand, usually a function of luck.

Goldratt has said on numerous occasions that there is no such thing as product cost: there are costs to maintain an operation at a given capacity level, but product cost is a mathematical invention created by accountants to overcome the limitation of employees making decision without complete information.[7] That being said, the majority of production costs do not vary directly with incremental production of a single unit. Additionally, most production costs are required to maintain a system of production, regardless of the number of units created.

Another problem that comes from using volume in the AUC calculation is the source of the number for volume. Suppose you're in charge of developing the product cost calculation for the upcoming quarter, which will be handed off to sales so that sales reps know what products to emphasize based on their gross margins. You'll need to get a number for volume. This number will most likely come from the sales forecast, which for most companies is not very accurate, especially at the unit level. This means that the number you'll calculate for product cost and gross margin is also not very accurate.

Another issue with AUC and its lack of accuracy comes from how capacity is obtained. As Exhibit 21.7 details, capacity is procured in blocks. A company is rarely able to buy a single production unit of capacity, even if they outsource to a third party. The stairstep line represents a company buying capacity in blocks from adding a new machine or adding an additional line or shift. The half-dome shape represents typical demand for a product over its life cycle. The exhibit shows that once a company installs new capacity, it spends a period of time with excess capacity versus demand. But as demand for products grow, the company will spend time capacity constrained until they obtain additional capacity. The only time the burdening of overhead to products exactly matches, is at the crossing points on the graph, which are few and far between.

In throughput accounting, the cost of the product is equal to its TVC. As long as the price you are getting for a product is greater than the TVC, there is a positive impact to contribution margin. This gives companies much more pricing flexibility.

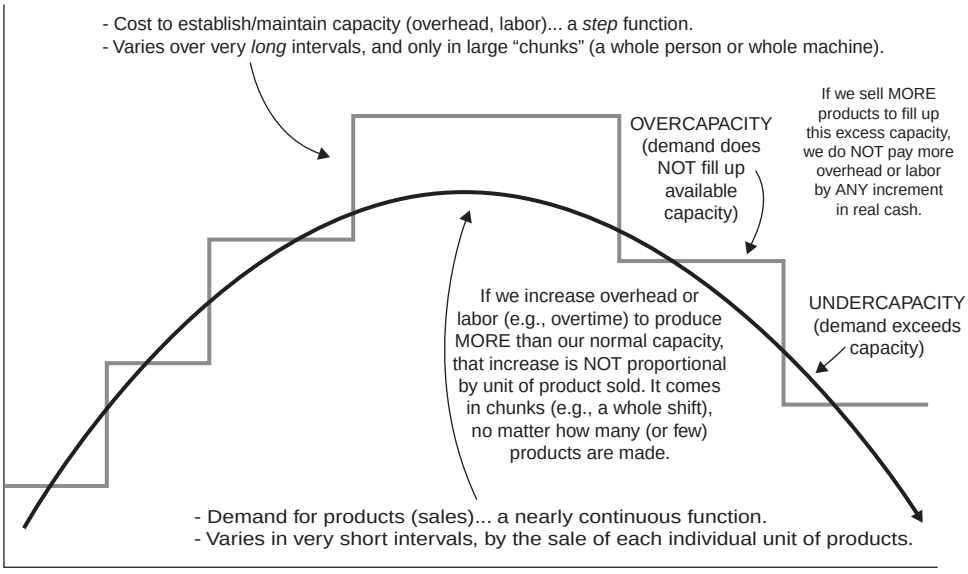


EXHIBIT 21.7 Capacity Constraints in Production
Source: Chart developed by H. William Dettmer, Goal Systems International, www.goalsys.com.

ANALYZING PRODUCTS BASED ON THROUGHPUT PER CONSTRAINT UNIT

One of the most powerful applications of throughput accounting is its ability to help companies shape demand toward those products that bring the most money to the bottom line. When companies do not have an internal constraint—in other words, the constraint is in the market—companies should make and sell all products that have demand and can be sold for a price higher than their TVC. That being the said, companies need to be aware of which products have the highest throughput and should encourage their salesforce to sell those products over products with lower throughput.

Once an internal constraint exists, the product emphasis for sales should shift to those products that have the highest throughput per constraint unit (T/CU).

For example, suppose a company manufactures cellular phones with different camera capabilities at varying megapixels (MP). The total demand for their cameras for the upcoming quarter is 1,200,0000 cell phones; however, the company has the ability to produce only 1 million cameras due to an internal constraint within their final testing machines. Each product must go through the constraint and has the financial profile shown in Exhibit 21.8.

Based on the throughput accounting analysis, sales would want to shape demand toward the 16MP phone.

To illustrate a very common scenario that one is likely to see when implementing T/CU, let's bring in the gross margin percentages. Gross margin should be shown only to illustrate the fact that AUC provides the wrong product emphasis (see Exhibit 21.9).

EXHIBIT 21.8 Sample Cell Phone Constraints with T/CU

Product	Price	TVC	Throughput	Constraint Time	T/CU
Cell phone with camera 6 MP	$43.00	$31.00	$12.00	11	$1.09
Cell phone with camera 8 MP	$65.00	$36.00	$29.00	30	$0.97
Cell phone with camera 16 MP	$68.00	$40.00	$28.00	25	$1.12
Cell phone with camera 32 MP	$87.00	$39.00	$48.00	100	$0.48

Because cost accounting assumes that all work stations are equal and does not consider the role of a constraint when generating product cost, the basis for gross margin, it is very common to see products with very low T/CU get emphasis over products with significantly higher T/CU.

To illustrate this point further for those that will inevitably struggle with the concept, it's helpful to take this example one step further.

The total available hours of the system constraint, final test is 11 million. If the company were able to make and sell each product line exclusively using all the available test hours of final test, the company would have the financial profile seen in Exhibit 21.10.

In other words, if the company made only 6MP phones using all the available test hours and sold all the phones it produced, it would make $12,320,000 in throughput this quarter. However, if the company made and sold only 32MP phones, the phones with the highest gross margin, the company would make only $5,280,000—a difference of $7,040,000.

Again, the reason for this is due to the fact that cost accounting does not properly account for the existence of a constraint. Companies need to understand that when they have an internal system constraint—a very common occurrence—they go from selling products to selling constraint units!

Companies still have to sell what customers are demanding, so you wouldn't make only 16MP phones if all the demand were in 32MP. However, companies have more power than they give themselves credit for when it comes to demand shaping, through pricing or sales programs. Throughput accounting enables companies to understand the true profitability of their products.

In the example, the constraint unit in the T/CU calculation was time on the constraint, but the constraint unit could have very easily been the motherboard supplier, in which case the calculation would be throughput per motherboard.

EXHIBIT 21.9 Sample Cell Phone Gross Margins

Product	Price	TVC	Throughput	Constraint Time	T/CU	Gross Margin
Cell phone with camera 6 MP	$43.00	$31.00	$12.00	11	$1.09	8%
Cell phone with camera 8 MP	$65.00	$36.00	$29.00	30	$0.97	20%
Cell phone with camera 16 MP	$68.00	$40.00	$28.00	25	$1.12	13%
Cell phone with camera 32 MP	$87.00	$39.00	$48.00	100	$0.48	32%

EXHIBIT 21.10 Sample Cell Phone Financial Profile

Product	Price	TVC	Through-put	Constraint Time	T/CU	Gross Margin	Volume	Through-put
Cell phone with camera 6MP	$43	$31	$12	11	$1.09	8%	1,000,000	$12,000,000
Cell phone with camera 8MP	$65	$36	$29	30	$0.97	20%	366,667	$10,633,333
Cell phone with camera 16MP	$68	$40	$28	25	$1.12	13%	440,000	$12,320,000
Cell phone with camera 32MP	$87	$39	$48	100	$0.48	32%	110,000	$5,280,000

HOW CAN A COMPANY INCREASE T/CU?

There are a number of things that a company can do to increase its products T/CU.

First, they can raise prices. However, for most companies this is a not an option. Second, they can reduce the time a product spends on the capacity constraint. This is a very viable option that produces enormous result. For example, if we were able to make a change in the processing of 32MP phones that would reduce the test time by 20 hours, for a total of 80 hour per phone, the T/CU would go from $0.48 to $0.60. Also, an additional 20 hours would be available per 32MP phone, which would be available toward the production of other MP phones.

Third, they can reduce TVC, which would increase the total throughput of the phone. Although this is a viable option, it is very important that companies keep a sharp eye on the impact at the system constraint. It is not uncommon for a company to reduce their raw material cost by going with a cheaper component, only to find that it increases the total processing time or negatively impacts yield on the system constraint.

Fourth, they can increase the yields at the CCR. It is critical for companies to make sure that only the highest quality material passes through the Capacity Constraint, because if a work-in-process product fails within or after processing by the constraint, the scrap or rework cost increases dramatically. For example, suppose the 16MP phone goes through the capacity constraint and fails, if it's decided that the product will go through a rework loop, it must pass through the capacity constraint again; hence, its total capacity constraint processing time goes from 25 hours to 50 hours, lowering its T/CU from $1.12 to $0.56. To fix this problem, it's common for companies to move their quality control from the end of their production line to the front of the capacity constraint, thereby ensuring that only good parts go through the constraint.

This also brings up the issue of scrap verses rework. If a work-in-process product is scrapped before it goes through the capacity constraint, the cost of scrap is equal to its TVC up to the point of scrap. However, if the product is scrapped after it has gone through the capacity constraint, the cost is significantly higher. By using throughput accounting companies can make better decisions regarding scrape versus rework.

KEY DECISIONS AREAS TO APPLY THROUGHPUT ACCOUNTING

The following are a handful of decisions that can be made more effectively with throughput accounting.

- *Product emphasis.* As mentioned earlier, by using throughput and T/CU companies get a much better picture of profitability of their products. As a result, companies should attempt to shape demand toward those products that bring in the most throughput to the bottom line. Two of the most common ways to shape demand are through pricing and/or by scaling down a company's product offering to those products that maximize throughput.
- *Scrap versus rework.* Companies can leverage throughput accounting to understand if it makes financial sense to put a product through a capacity constraint twice or scrap the product.
- *Outsourcing decisions.* T, I, and OE analysis should be used to evaluate how much capacity and net income will be made resulting from outsourcing. When doing an outsourcing analyst, make sure to properly account for material flow risk, which is typically compensated with additional inventory.
- *Product addition.* It's critical that companies use T, I, and OE analysis to understand what impact new products will have to the system constraint. The beauty of throughput accounting is that it focuses design efforts at the correct location to maximize net income, the system constraint.
- *Product transitions.* Companies that use cost accounting for decision support typically make the mistake of killing a *cash cow* in the middle of their life cycle.[8]

  For example, it is very common for companies using cost accounting for decision support to make the wrong decision when faced with the following, as shown in Exhibit 21.11.

  The 6MP camera is out in the market with strong demand of 1 million units for the upcoming quarter. However, due to the gross margin differences, there is tremendous pressure on sales to sell the 8MP camera, which has a very low forecasted demand of only 150,000. Facing this scenario, companies are tempted to take the *build it and go sell it* approach. However, a company that uses throughput accounting would understand that the 6MP camera is a *cash cow* through their system constraint and would try to keep sales of the 6MP camera going for as long as possible. Clearly, there is time to market advantages that could come from being first to the market with the 8MP camera; however,

EXHIBIT 21.11 Sample Cell Phone Company Using Cost Accounting

Product	Price	TVC	Through-put	Constraint Time	T/CU	Demand	Gross Margin
Cell phone with camera 6MP	$43	$31	$12	11	$1.09	1,000,000	8%
Cell phone with camera 8MP	$65	$36	$29	60	$0.48	150,000	35%

the point here is to go into the decision with the correct analysis, not the analysis that is giving the wrong answer. The optimal time to release the 8MP camera is when the T/CU is at or approaching the T/CU of the transitioned product.

- *Capital investment.* This decision goes back to throughput accounting's definition of return on investment.

$$\text{ROI} = (\Delta\text{T} - \Delta\text{OE})/\Delta\text{I} \tag{21.4}$$

The ROI equation allows companies to accurately measure impact to capacity and net profit resulting from capital investments.

- *Process improvement expenditures.* The use of T, I, and OE makes it very easy for companies to measure the benefit from improving a process. Those process improvement efforts that focus on the system constraint will have a positive impact to throughput and should demand attention from Six Sigma, Lean, and other improvement efforts.
- *Marketing potential.* Comparing markets using T, I, and OE analysis helps companies decide which markets have the potential to produce the most throughput.
- *Staffing decision.* T, I, and OE analysis here will help companies understand how much additional capacity they will get by adding additional staff or shifts.
- *Project selection.* The use of T, I, and OE analysis in project selection is something that can show immediate impact. One of the shortfalls of Six Sigma and Lean is that it generates a handful of projects that minimize cost as the primary success driver. By using throughput accounting, companies prioritize their resources to those projects that drive the highest throughput while simultaneously reducing investment and operating expense. Dr. Goldratt has said at numerous *Theory of Constraints International Certification Organization* Conferences that once companies use the Theory of Constraint to guide Six Sigma and LEAN efforts through the use of the "Five Focusing Steps," the benefit will be tremendous.[9]

SUMMARY

Constraint management and throughput accounting have tremendous benefits for companies currently leveraging cost accounting for decision making and risk management. Although cost accounting must be used for reporting and taxes, there is no rule in place that mandates its use for decision making. As a matter of fact, it should now be very clear that the use of cost accounting for decision making commonly leads to suboptimal result. Although the examples in this chapter focused on production companies, the application of constraint management as well as throughput accounting can be very easily applied to service-based companies.[10]

Only by taking into account the impact of a decision to a company's system constraint can a company be sure that its decisions are taking a holistic approach—balancing opportunities and risk. Remember, constraints are not bad, they just are. Throughput accounting will enable you to take control of them and maximize your goal.

APPENDIX: COMMON QUESTIONS AND ANSWERS

1. Do we need a new accounting system to run throughput accounting?

   No, throughput accounting is not an accounting process/system. It is a decision support tool and does not replace standard accounting practices. Additionally, all of the numbers that we need to calculate T, I, OE, TVC, and T/CU typically exist within companies existing accounting systems.

2. Why use TVC and not AUC?

   Because TVC does not change with volume, hence it cannot be manipulated. Unless raw material cost are reduced as a result of economies of scale, TVC for a product will be the same whether you produce one or ten thousand. Also, TVC includes only those costs that vary directly with the number of units sold, thus eliminating the distortion caused by product cost allocation.

3. Why can't we just increase the depreciation burden for the product based on its constraint usage?

   Since the constraint dictates how much money a company is capable of making, the burden allocation a company would apply to the product using standard unit costing would still be significantly low. Additionally, the number would still be able to be manipulated through volume changes.

4. How can we be sure that we're covering the cost of our investments?

   If you use ROI as defined by throughput accounting ($T - \Delta OE/\Delta I$), you have already justified the investment based on the increase in net profit that will result from the investment; thus, you don't have to force utilization to justify a purchase you've already made.

5. Does throughput accounting consider multiple constraints?

   Constraint management, the process throughput accounting is based on, looks at primary control points in a system. If at some point in time there are multiple constraints (anything that has more demand than capacity), throughput accounting accounts for that situation.

6. Does this mean we'll produce only those drives that have the highest T/CU?

   No, we still must respond to customer demand and fulfill our strategic commitments. However, throughput accounting will tell us which drives have the highest net contribution per constraint unit. It's up to us to decide what to do with that information.

7. Why isn't labor considered in the calculation for truly variable cost?

   Most companies pay employee wages as a function of time (hourly) not by the piece past. As a result, the number of products a single employee can produce in an hour is not proportional one to one with a single unit of production. For example, in an hour an employee could make say six units if they're having a good day, two units if they're having a bad day, but on average four units a

day. Since labor is not proportional to a single unit of production, throughput accounting considers labor to be an operating expense.

8. Our constraint moves all the time, why aren't we constantly changing the throughput-per-constraint-unit calculation?

   Companies need to distinguish between a bottleneck and a system/strategic constraint. A bottleneck is a temporary phenomenon that can be overcome in less than a quarter, versus a system/strategic constraint, which truly dictates the capacity of the operation. In other words, once a bottleneck is overcome, where in the system does the choke point always fall back to?

NOTES

1. Eliyahu M. Goldratt and Jeff Cox, *The Goal: A Process of Ongoing Improvement* (Great Barrington, MA: North River Press, 1984).
2. For more information on the Five Focusing Steps of Constraint Management, see Goldratt and Cox, 1984.
3. See note 1.
4. For articles and white papers, see www.eligoldratt.com. Suggested books include: Thomas Corbett, *Throughput Accounting* (Great Barrington, MA: North River Press, 1998), John A. Caspari and Pamela Caspari, *Management Dynamics: Merging Constraints Accounting to Drive Improvement* (Hoboken, NJ: John Wiley & Sons, 2004); and Steven M. Bragg, *Throughput Accounting: A Guide to Constraint Management* (Hoboken, NJ: John Wiley & Sons, 2007).
5. For more detailed examples, see Caspari and Caspari, 2004.
6. For more information on the TOC thinking process, see H. William Dettmer, *The Logical Thinking Process: A Systems Approach to Complex Problem Solving* (Milwaukee, WI: ASQ Press, 2007).
7. *Beyond the Goal: Eliyahu Goldratt Speaks on the Theory of Constraints* (Your Coach in a Box audio book CD).
8. *Cash cow* is a term that comes from the Boston Consulting Matrix, which is a chart created by Bruce Henderson for the Boston Consulting Group in 1970 to help corporations with analyzing their business units or product lines.
9. For more information on the Theory of Constraints International Certification Organization (TOC-ICO), see www.tocico.org.
10. For more information on applying TOC to service based companies, see John Arthur Ricketts, *Reaching the Goal: How Managers Improve a Services Business Using Goldratt's Theory of Constraints* (Armonk, NY: IBM Press, 2007).

CHAPTER 22

Environmental Consistency Confidence: Scientific Method in Financial Risk Management

Michael Mainelli, Ph.D.

INTRODUCTION

The application of the scientific paradigm to business operations transformed management thinking in the early part of the twentieth century. A plethora of management theorizing since often obscures the simplicity at the core of the scientific paradigm. One approach, environmental consistency confidence, restores statistical correlation to its rightful place at the core of financial risk management. For financial services organizations, statistical correlation integrates well with existing key risk indicator (KRI) initiatives. Through environmental consistency confidence, financial organizations understand the limits of their environmental comprehension.

In late nineteenth century, Frederick Winslow Taylor promoted scientific management. The legacy of Taylor's early attempts to systematize management and processes through rigorous observation and experimentation led to the quality control movement of the 1920s, operations research and cybernetics of the 1940s, and Total Quality Management (TQM) of the 1980s, leading through to today's Six Sigma and Lean manufacturing. The aim of scientific management is to produce knowledge that improves organizations using the scientific method. Taylor promoted scientific management for all work, such as the management of universities or government.

The scientific method is based on the assumption that reasoning about experiences creates knowledge. Aristotle set out a threefold scheme of abductive, deductive, and inductive reasoning. Inductive reasoning generalizes from a limited set of

I would like to thank Adrian Berendt, Brandon Davies, Christopher Hall, Ian Harris, Matthew Leitch, Jan-Peter Onstwedder, Jürgen Sehnert, Jürgen Strohhecker, and Justin Wilson for helping to develop some of the thinking behind this article, though not to claim they agree with all of it.

This chapter was adapted from an earlier version by Michael Mainelli ("Correlation Causes Questions: Environmental Consistency Confidence in Wholesale Financial Institutions," in *Frontiers of Risk Management,* edited by Dennis Cox, 94–100. Euromoney Books, 2007).

observations—from the particular to the general—"every swan we've seen so far is white, so all swans must be white." Deductive reasoning moves from a set of propositions to a conclusion—from the general to the particular—"all swans are white; this bird is a swan; this bird is white." But neither inductive nor deductive reasoning is creative. Abductive reasoning is creative, generating a set of hypotheses and choosing the one that, if true, best explains the observation "If a bird is white, perhaps it's related to other white birds we've previously called 'swans,' or perhaps it's been painted white by the nearby paint factory"—from observations to theories. Abductive reasoning prefers one theory based on some criteria, often parsimony in explanation, such as Occam's razor: "All other things being equal, the simplest explanation is the best."

The scientific method is the application of a process to the creation of knowledge from experience. The hypothetico-deductive model is perhaps the most common description of the scientific method, algorithmically expressed as:

1. "Gather data (observations about something that is unknown, unexplained, or new);
2. Hypothesize an explanation for those observations;
3. Deduce a consequence of that explanation (a prediction);
4. Formulate an experiment to see if the predicted consequence is observed;
5. Wait for corroboration. If there is corroboration, go to step 3. If not, the hypothesis is falsified. Go to step 2."[1]

The scientific method is hardly a sausage machine. William Whewell noted in the nineteenth century that that "invention, sagacity, genius" are required at every step in scientific method. Perhaps the most interesting twentieth-century insight into the scientific method came from Karl Popper, who asserted that a hypothesis, proposition, or theory is scientific only if it is falsifiable. Popper's assertion challenges the idea of eternal truths because only by providing a means for its own falsification can a scientific theory be considered a valid theory. Every scientific theory must provide the means of its own destruction and thus is temporary or transient, never an immutable law.

Most managers would consider at least a part of their management style to be "scientific." They deal with numbers. They use numbers to spot anomalies, examine them for further evidence, and make decisions based, at least partly, on numerical reasoning processes. MBAs graduate having studied "quant" skills. Accountants deploy their arithmetical techniques across businesses. It is true that Aristotle's inductive reasoning process—"every Christmas our sales go up, thus our sales will go up this Christmas"—and abductive reasoning process—"our sales go up at Christmas because people like to give presents, or because people buy our fuel oil"—are widely used. However, the deductive formality of the scientific process in management is rarely applied.

1. Gather data (observations about something that is unknown, unexplained, or new)—"we've detected an unusual decrease in trade closure times."
2. Hypothesize an explanation for those observations—"using our abductive methods we can surmise that 'clients have changed their purchasing behavior' or 'our new computer system has sped things up' or 'our traders are up to something.'"

3. Deduce a consequence of that explanation (a prediction)—"we should see 'a change in phone call lengths' or 'contrasted times between manual and computer trades should be larger' or 'the nature of trades have changed.'"
4. Formulate an experiment to see if the predicted consequence is observed—don't accept an easy explanation, go and check.
5. Wait for corroboration. If there is corroboration, go to step 3. If not, the hypothesis is falsified. Go to step 2.

PARADIGMS APPLIED—VALUES, CONTROL, REENGINEERING, AND COSTING

When it comes to risk management, the core process should be one of scientific management. Wholesale financial institutions frequently lack financial risk management structured for its own sake, rather than as a response to regulatory pressures. Wholesale financial institutions tend to respond positively to regulatory initiatives, but otherwise do what everyone else is doing. Wholesale financial institutions have deployed at least four generic approaches for managing and modeling operational risk, with limited success—shared values, control structures, reengineering, and costing risk.

Let's start with "shared values" approaches. While not denying the importance of culture[2]—"would one rather have a bunch of honest people in a loose system or a bunch of crooks in a tight system?"—and its crucial role as the starting point for risk management, cultural change is hard to formalize. At one extreme, one can parody culturally based risk programs as "rah! rah!" cheerleading—"every day in every way, let's reduce risk"—but people in organizations do need to share values on risk awareness, assessment, and action. Shared values are essential but insufficient for financial risk management.

Another common approach is "control structures." Often denigrated as "tick bashing," control structure approaches are particularly common in regulated industries. The difficulties with control structures are legion, for example, tough to design, often full of contradictions (Catch-22s), difficult to roll back, expensive to change. Control structures often result in a command-and-control organization, rather than a commercial one, with costs frequently exceeding not just the potential benefits but also the available time.[3] Some institutions deploy risk dashboards or "radars"—tools that aggregate procedural compliance—with little consideration of the human systems within which this approach is being applied. While this heuristic approach is culturally suited to banks (bureaucratic "tick bashing" and form filling with which they are familiar), excessive control structures undermine and contradict shared values.

The positive view of undermining is "working the system," but the negative view is lying. A simple example, managers inculcate a lying culture among subordinates to avoid chain-of-command pressures on targets—"I know you can't lock the computer door on our African computer center because you've been awaiting air conditioner repair for the past five days, but could you just tick the box so my boss stops asking about it on his summary risk report?" Another example: people repeatedly answer questions with the desired answer; for example, "Does this deal have any legal issues?" strongly suggested for an easy life answer "no," thus penalizing honest

thinking. Finally, the resulting RAG (red-amber-green) reports cannot be readily summarized or contrasted—five open computer room door incidents may be rated more important than a single total power outage.

Reengineering via process modeling and redesign is used in many industries, including finance. Many financial institutions document their operations in order to analyze their operational risks. Many of the tools used to document operations are the same tools used by system dynamics simulation models. This happy coincidence led many institutions to experiment with system dynamics techniques, but then they encountered problems of validating the models and chaos theory effects, that is, extreme sensitivity to initial conditions, as well as the expense of trying to maintain models of business operations in a fast-changing environment. Reengineering is a good tool for improving processes, but does not sit at the heart of risk management.

A risk management approach does integrate with financial and economic theory when "costing risk," typically using economic cost of capital and value at risk (VAR). The basic idea is to build a large, stochastic model of risks and use Monte Carlo simulations to calculate a VAR that allows a financial institution to set aside an appropriate amount of capital—economic cost of capital—per division or product line. This approach requires probability distributions of operational risk, market movements, and credit defaults. Yes, it is difficult and at an early stage, but this approach has merit both for management and regulators. However, it does not provide a core scientific management process.

ENVIRONMENTAL CONSISTENCY CONFIDENCE—STATISTICAL HEAD, CULTURAL HEART

What distinguishes good financial risk management from bad? In a nutshell, it's a scientific approach to risk. At the core of the scientific approach is a statistical engine room of some form:

> *Statistical and applied probabilistic knowledge is the core of knowledge; statistics is what tells you if something is true, false, or merely anecdotal; it is the "logic of science"; it is the instrument of risk-taking; it is the applied tools of epistemology.*[4]

Environmental consistency confidence is an approach to risk management that says, "If you can predict incidents and losses with some degree of confidence, then you have some ability to manage your risks." You are confident to some degree that outcomes are consistent with your environment and your activities. The converse—if you can't predict your incidents and losses—implies either that things are completely random and thus there is no need for management, or that you're collecting the wrong data. Knowing that incidents and losses are predictable then leads to application of the scientific paradigm. From a proven hypothesis, financial risk tools such as culture change, controls, process reengineering, or risk costing can be usefully applied.

A few years ago, when promoting environmental consistency confidence to one trading firm, Z/Yen posed a tough question: "Why can't you predict the losses and incidents flowing from today's trading?" The idea was to look at the environmental and activity statistics for each day and use multivariate statistics to see how strong the

correlation was with incidents and losses flowing from that day. It is often said that "correlation doesn't demonstrate causation." That is true, but "correlation should cause questions."

> *For example, you might be wondering why people make mistakes when they enter data into a particular system. Some people make more mistakes than others. Are they careless? Do they need training? Is the system hard to use? Do people have too much work to do? Do they make more errors when they work on into the evening? Is there something about the particular data they enter that makes errors more likely? Do changes to the entry screen increase or decrease errors?*[5]

The core of environmental consistency confidence is using modern statistical models to manage financial institutions through the examination of correlations between activity and outcomes. Environmental consistency confidence starts with the idea that the organization is a large black box. If the outputs of the box can be predicted from the inputs using multivariate statistics, then the scientific management process can be deployed, abductively (creatively), inductively (experience), and deductively (analytically). The key elements of Environmental Consistency Confidence are:

- A strong database of day-to-day environmental factors and trading activities.
- A database of incidents and losses (or errors or nonconformities or other measures of poor performance).
- A unit tasked with predicting future incidents and losses from current factors and activities.
- A "confidence" measure (typically R^2) from the unit is about predictive accuracy.

If the unit is highly confident of predictions, then management has work to do, typically deploying scientific management techniques. If the unit is unsure, less confident, then more and better data or predictive techniques need to be sought. Overall, when inputs from the environment and the activity levels match overall outputs, then the organization is "consistent" with its environment. The idea is not just to amass facts, but to turn anomalies and prediction variances into science:

> *Science is facts. Just as houses are made of stones, so is science made of facts. But a pile of stones is not a house and a collection of facts is not necessarily science.*[6]

WHAT IS A KEY RISK INDICATOR (KRI)?

Frequently, predicting losses and incidents revolves around correlations with key risk indicators (KRIs). A working definition for KRIs is "regular measurement data that indicates the risk profile of particular activities." KRIs help to form an input for economic capital calculations by producing estimates of future operational risk losses and thus helping to set a base level of capital for operational risk. KRIs

can be environmental, operational, or financial. KRIs are increasingly important to regulators.

> *Key Risk Indicators: risk indicators are statistics and/or metrics, often financial, which can provide insight into a bank's risk position. These indicators should be reviewed on a periodic basis (often monthly or quarterly) to alert banks to changes that may be indicative of risk concerns.*[7]

For wholesale financial institutions, environmental consistency confidence is strongly linked with predictive key risk indicators for losses and incidents (PKRI ⇨ LI). The important point to note is that people can suggest many possible risk indicators (RIs), but they are not KRIs unless they are shown to have predictive capability for estimating losses and incidents. A KRI must contribute to the predictability of losses and incidents in order to be validated as a KRI. If an RI does not predict losses or incidents, it remains an interesting hypothesis, someone's unvalidated opinion. The scientific approach to managing risk using statistics also involves trying to discover what the indicators *should* have been. In other words, what drives operational risk? We describe this approach as predictive key risk indicators to/from loss/incidents prediction (PKRI ⇔ LI).

Experience does help to identify the true drivers of operational risk and should help focus attention and control actions, but the PKRI ⇔ LI approach supports and validates (or invalidates) expert judgment of true drivers of operational risk losses. The intention of this approach is not to replace expert judgment, but to support that judgment in a more scientific way in an ever-changing environment. For instance, environmental indicators (that might turn out to be KRIs) could be such things as trading volumes and volatilities on major commodities or foreign exchange markets. Operational indicators (that might be KRIs) could be general activity levels in the business, numbers of deals, mix of deals, failed trades, number of amendments, reporting speed, staff turnover, overtime, or information technology (IT) downtime. Financial indicators (that might be KRIs) could be things such as deal volatility, dealing profit, activity-based costing variances, or value of amendments.

In a sense, the choice is between what is currently done informally (no significant business lacks RIs) and what could be done better through more formality, statistics, and science to make them KRIs. For each KRI, there needs to be definition and specification. Exhibit 22.1 sets out the characteristics of a KRI as seen by the Risk Management Association.

CASE STUDY: GLOBAL COMMODITIES FIRM

A large global commodities firm active not only in a number of commodity markets but also foreign exchange and fixed income piloted the PKRI ⇔ LI approach in one large trading unit. Overall, as might be expected, the findings were that low-volume and low-complexity days in a low- or high-stress environment were fine. Intriguingly, low volume and low complexity were slightly worse in a low-stress environment. High-volume and high-complexity days in either a low- or high-stress environment indicated relatively high forthcoming losses and incidents. High-volume and low-complexity days caused relatively few difficulties. Low-volume and

EXHIBIT 22.1 Characteristics of Key Risk Indicators

Effectiveness	Comparability	Ease of Use
Indicators should...	Indicators should...	Indicators should...
1. Apply to at least one risk point, one specific risk category, and one business function. 2. Be measurable at specific points in time. 3. Reflect objective measurement rather than subjective judgment. 4. Track at least one aspect of the loss profile or event history, such as frequency, average severity, cumulative loss, or near-miss rates. 5. Provide useful management information.	1. Be quantified as an amount, a percentage, or a ratio. 2. Be reasonably precise and define quantity. 3. Have values that are comparable over time. 4. Be comparable internally across businesses. 5. Be reported with primary values and be meaningful without interpretation to some more subjective measure. 6. Be auditable. 7. Be identified as comparable across organizations (if in fact they are).	1. Be available reliably on a timely basis. 2. Be cost effective to collect. 3. Be readily understood and communicated.

high-complexity days were fine in a low-stress environment, but poor in a high-stress environment. The key control point going forward was to make trading complex products harder in high-stress or high-volume situations.

While the predictive success was adequate only in the pilot, with an R^2 approaching 0.5, the approach was seen to have merit and the firm rolled out the PKRI ⇔ LI methodology globally across several business units. It was telling that the PKRI ⇔ LI approach helped the commodities firm realize the importance of good data collection and use, and to identify areas where data specification, collection, validation, and integration could be markedly improved. Multivariate statistics, such as the use of support vector machines, did not add much value in the early stages; many of the predictive relationships were straightforward, for example, large numbers of deal amendments lead to later reconciliation problems.

One example of the scientific method being applied to a business problem was in trader training. Trading managers felt that job training was useless, but were afraid to say so in front of the human resources team. For trading managers, people either "had it, or they didn't." For human resources, almost any process problems required "more training." The hypothesis was "increased training leads to fewer errors." The PKRI ⇔ LI approach was to see if low training was predictive of losses

and incidents. In the event, the answer was "no." Losses and incidents were fairly random for the first six months of trader employment. Apparently, team leaders weed out poor traders within the first six months of employment. From this point on, until approximately four or five years of trading have passed, training does not correlate with reduced losses or incidents, just poorer profit performance due to days lost in training. After four or five years of employment losses and incidents begin to rise, presumably traders "burn out." But again, increased training made no difference. Human resources countered that perhaps it was the "wrong kind of training." Perhaps, but the trading managers wanted experimental tests of efficacy before rolling out new, costly training programs/scientific management.

PREDICTIVE KEY RISK INDICATORS FOR LOSSES AND INCIDENTS (PKRI ⇔ LI) ISSUES

There is overlap between KRIs and key performance indicators (KPIs). It would be easy to say that KRIs are forward looking and KPIs are backward looking, but far too simplistic. For instance, high trading volumes and high volatility on one day might be good performance indicators predicting a high likelihood of good future financial performance turnout for that day, but also indicative of emerging operational risks from that day. A KRI such as the number of lawsuits received by a particular function might change very little for long periods. In this case one might wish to examine "lawsuits in period" or "estimated settlement values" or other more sensitive measures than just a very slow-changing "outstanding lawsuits." However, what matters is whether the KRI contributes to the capability of predicting operational losses/incidents, not its variability.

KRIs that increase in some ranges and decrease in others can cause confusion as KRIs are not necessarily linear. For example, staff overtime might be an example of a KRI with a bell-shaped curve. No overtime may indicate some level of risk, as people aren't paying attention or do tasks too infrequently; modest levels of overtime may indicate less risk as staff are now doing a lot of familiar tasks; and high rates of overtime may indicate increased risk again through stress. KRIs help to set ranges of acceptable activity levels. There can be step changes in operational risk associated with a KRI. For instance, a handful of outstanding orders at the close of day may be normal, but risk might increase markedly when there are over a dozen outstanding orders. KRIs should vary as risk changes, but they don't have to vary linearly.

CASE STUDY: EUROPEAN INVESTMENT BANK

One European investment bank used three years of data to predict losses/incidents such as deal problems, IT downtime, and staff turnover over a six-month period. It achieved reasonable predictive success using multivariate statistical techniques such as support vector machines, with R^2 approaching 0.9 at times, though more frequently 0.6 (i.e., 60 percent of losses can be predicted). Exhibit 22.2 shows a high-level snippet, giving a flavor of the data.

Note that some of the items in this snippet (e.g., HR joiners/leavers or IT disruption at the system level) can in practice be very hard to obtain. It was also noteworthy

EXHIBIT 22.2 European Investment Bank's Data Sample

Location ID	HR-Head-count #	HR-Joiners in month	HR-Leavers in month	IT-System Disrup-tion Incidents	IT-System Down-time	FO-Trade Volume #	FO-Trade Amend-ments #	OPS-Nostro Breaks #	OPS-Stock Breaks #	OPS-Intersys-tem Breaks #	OPS-Failed Trades #	OPS-Unmatched Trades #	RIS-Market Risk Limit Breaches #	AU-High Risk O/S Overdue Audit Issues #	AU-High Risk O/S Audit Issues #
1	136	6	11	2	35:07	19218	317.1	3	9	6	463	52	0	0	4.5
2	121	6	11	2	3:13	8999	0	17	4	2	26	0	3	0	4.5
3	23	6	11	0	0	661	8.7	3	0	0	0	7	0	0	4.5
4	30	6	11	0	0	4307	80.5	7	1	1	17	0	1	0	4.5
...															
n															

that, as a data-driven approach, PKRI ⇔ LI projects are only as good as the data put into them—"garbage in, garbage out." In some areas, the data may not be at all predictive. Data quality can vary over time in hard-to-spot ways and interact with wider systems, particularly the people in the systems. For instance, in this trial of PKRI ⇔ LI, the IT department was upset at IT downtime being considered a "key risk indicator" and unilaterally changed the KRI to "unplanned" IT downtime, skewing the predicted losses. This change was spotted when using the dynamic anomaly and pattern response (DAPR) system to run the reverse LI ⇔ PKRI prediction as a quality control—another example of Goodhart's Law, "when a measure becomes a target, it ceases to be a good measure" (as restated by Professor Marilyn Strathern).

So what about all the key risk indicators that have not been taken seriously, such as water and electrical utilities? They appear to be important when considering KRIs in developing world locations, but are less critical in the developed world. In major financial centers such as New York and London, many business continuity risks are taken for granted—for example, an absence of natural threats such as hurricanes or flooding, yet London has a history of significant flooding from the Thames Barrier. Of course, 9/11 established the vulnerability of New York to major disruptions. Seismic issues such as earthquakes and tsunamis, or health issues such as the severe acute respiratory syndrome (SARS) pandemic don't seem to feature. There are also numerous personal issues that don't feature—such as schools, opening bank accounts, work permits, arranging for utilities, personal safety—any of which could scupper the trading floor. Understandably, people will care about events and issues that they are conscious of. There are many issues that could have us looking back many years from now and sadly remising about how trading ceased to function when infectious diseases became too dangerous to have people so highly concentrated, or when people wanted to avoid concentrating in a given area because of the risk of terrorism. The PKRI ⇔ LI approach is one for regular management, not extreme events.

Adrian Berendt points out that in operational risk there is a focus on the known and the known-unknown (those matters that we can comprehend and do something about), at the expense of the unknown-unknown. This may be why businesses focus more on disaster recovery from infrastructure disruption rather than climate change, even if the latter were a larger risk. Events may be viewed as serious if they knock out our single building, but (perversely) less serious if they knock out a whole city. The reasoning: "If there is a catastrophe, our customers will understand that we cannot service them, whereas, if we have had a computer glitch, they will go to our competitors." On global risks, "if the catastrophe is so great that all competitors are affected, nothing we do will make a difference and we won't be disproportionately disadvantage competitively."

Taleb cautions us on the limits of statistical methods.[8] Where the distributions are "thin-tailed," that is, close to normal and not seriously skewed, then statistical techniques work well. There are few "black swans" (i.e., more rare events than expected based on historical frequency). Where the payoffs are simple, then approaches such as environmental consistency confidence work well. However, where distribution tails are "fat" and payoffs complex, then statistical methods are fragile and susceptible to rare events. Unfortunately, the abundance of rare events, largely due to the fact that financial markets feed forward from interconnected human behaviors, leads people to impugn statistical techniques in finance. Some esoteric

EXHIBIT 22.3 DMAIC—Existing Product/Process/Service

Stage	Objectives
Define	Define the project goals and customer (internal and external) deliverables.
Measure	Measure the process to determine current performance.
Analyze	Analyze and determine the root cause(s) of the defects.
Improve	Improve the process by eliminating defects.
Control	Control future process performance.

risk calculations need extremely long time period data to prove that some extreme value calculation is true. If you can get quality data, sometimes it's provable. More often, when real-world data fails to fit, you find that you've used the wrong distribution or poor-quality data, too late. Rather than abandon all statistics because some extreme cases are problematic, the suggestion should be that people develop extremely strong environmental consistency confidence units, but be clear of their limitations. Other approaches, such as scenario planning or aggregated human judgment, may assist in evaluating rare, complex payoff situations. The frequency of black swan events argues for higher provisioning and increased redundancy regardless of some core numbers that environmental consistency confidence on its own might imply.

WHAT IS CURRENT PRACTICE?

Scientific management of wholesale financial operations is increasing. Managers in many investment banks (e.g., Bank of America, JPMorgan Chase) have publicly announced their pursuit of Six Sigma or their adherence to DMAIC (define, measure, analyze, improve, and control; see Exhibit 22.3) or DMADV (define, measure, analyze, design, and verify) Six Sigma (see Exhibit 22.4) approaches (originally from GE) when they have losses/incidents that they want to eliminate by eliminating root causes.

Six Sigma is clearly related to a dynamic system view of the organization, a cycle of tested feed-forward and feedback. This had led to greater interest in using predictive analytics in operational systems management. Several leading investment banks, using Six Sigma programs and statistical prediction techniques (predicting

EXHIBIT 22.4 DMADV—New Product/Process/Service

Stage	Objectives
Define	Define the project goals and customer (internal and external) deliverables.
Measure	Measure and determine customer needs and specifications.
Analyze	Analyze the process options to meet the customer needs.
Design	Design (detailed) the process to meet the customer needs.
Verify	Verify the design performance and ability to meet customer needs.

trades likely to need manual intervention), have managed to reduce trade failure rates from 8 percent to well below 4 percent over three years for vanilla products. As the cost per trade for trades requiring manual intervention can be up to 250 times more expensive than trades with straight-through-processing transaction, this is a very important cost-reduction mechanism, as well as resulting in a consequent large reduction in operational risk.

Another approach used in investment banking is predictive analytics. Predictive analytics feature where investment banks move towards automated filtering and detection of anomalies (DAPR).[9] Cruz notes that a number of banks are using DAPR approaches not just in compliance, but also as operational risk filters that collect "every cancellation or alteration made to a transaction or any differences between the attributes of a transaction in one system compared with another system. . . . Also, abnormal inputs (e.g., a lower volatility in a derivative) can be flagged and investigated. The filter will calculate the operational risk loss event and several other impacts on the organization."[10] See Exhibit 22.5.

Given the impact of the credit crunch on trust in all modeling, there is a tendency to assume that all statistical techniques are suspect. Statistical techniques tended to be focused solely on pricing as opposed to operational risk and overall systemic performance. In fact, many of the environmental consistency confidence techniques would have driven financial services people to pay much more attention to costing liquidity risk years ago.[11]

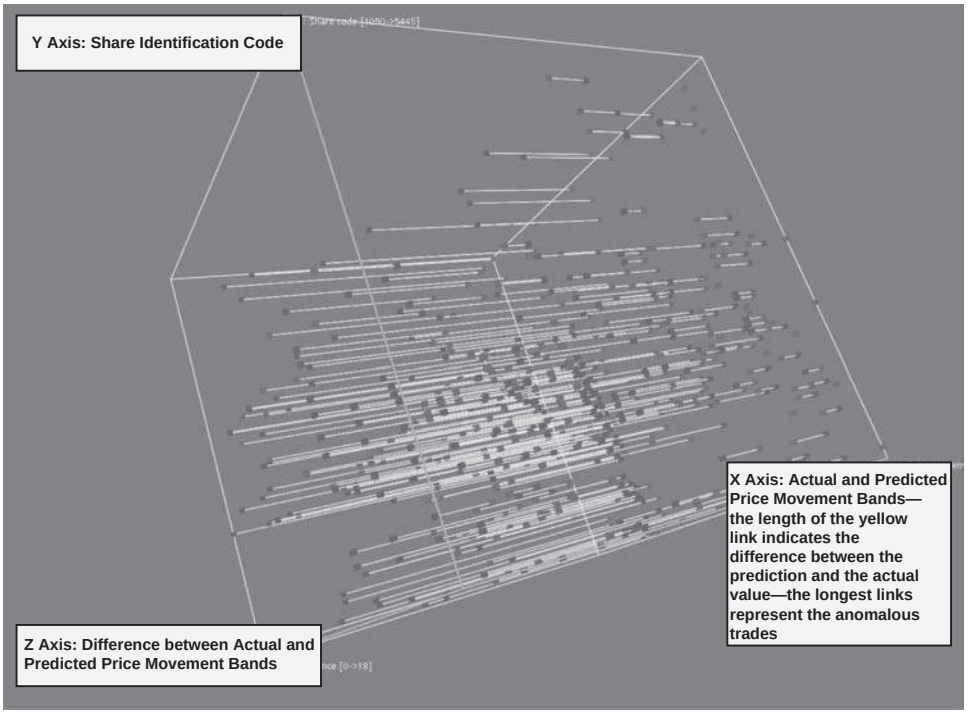


EXHIBIT 22.5 DAPR Support Vector Machine Example: Contrasting a Subset of Actual vs. Predicted Trade Price Bands

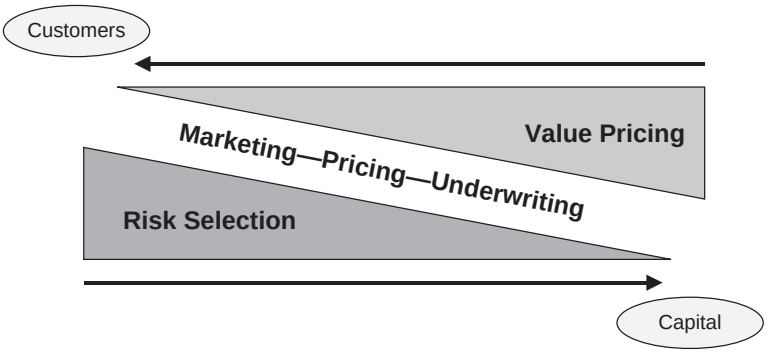


EXHIBIT 22.6 Value to Customers and Cost of Capital

BIGGER CANVASES FOR SCIENTIFIC MANAGEMENT

Successful KRIs are made up of a combination of factors, and not just from a single factor. In the opening line of *Anna Karenina,* Tolstoy provides a principle applicable to KRIs: "Happy families are all alike; every unhappy family is unhappy in its own way." Jared Diamond contends this principle describes situations in which a number of activities need to be executed correctly to achieve success, while failure can result from only one poorly executed activity. This applies to KRIs in which the evolving set of KRIs is essential, not just a single one in a single time frame, nor too many KRIs at the same time. This underscores the importance of multivariate statistics in any real-world use of KRIs.

We must not lose sight of the scientific method—we propose a hypothesis in which certain combinations of KRIs can help to predict future incidents and losses. We can test this hypothesis by examining a set of these combinations using available statistical tools. If the environmental activities and factors we apply are consistent with the outcomes, then we can have a higher degree of confidence that we are tracking the correct factors. Once we establish that we are tracking the correct factors, we can develop solutions or projects that mitigate or eliminate the causes. If our predictions fail, we have not tracked the correct factors. In such cases, we need to continue to quickly explore other factors as this suggests that our process or events may be out of control.

When we examine the broader wholesale finance system, we will discover related high-level systems that can also be predicted, not the risks. Exhibit 22.6 shows a relatively simple finance model in which risks are selected through marketing and positioning. The risks are then priced by ascertaining or at least attempting to ascertain the cost of capital and the difference in value to customers.

We can apply this finance abstract model to a KRI system as follows: for underwriting/trading, are we able to forecast losses and incidents; for sales and marketing, are we able to forecast sales; for pricing, are we able to forecast profitability? A KRI system has the following components:

- *Governance*. Use the organization's business objectives to create the operational risk framework definition. Next, calculate economic capital. Finally, establish a basic set of essential KRIs.

- *Input.* Start by winning the commitment of stakeholders. Next, assemble the needed resources. Finally, appoint a team to create the potential KRIs.
- *Process.* Start by supporting the efforts of operational risk managers such as collecting data, statistical testing and validating using statistics, making correlations and multivariate predictions, conducting cross-project discussions and training, and developing templates and standardized methodologies.
- *Output.* Start with the evaluation of KRIs. Next, focus on what Six Sigma calls the voice of the customer—both internal and external to determine. Finally, determine how this helps to manage the business better, so that your resources learn from their successes and their failures.
- *Monitoring.* Start by providing business process owners information up to directors, over to customers, down to and across project managers so that they are aligned and coordinated in their activities. Next, utilize *feedback* from KRI outcomes and the subsequent *feed-forward* inputs as part of the monitoring process. This provides new KRI ideas and helps to replan the KRI portfolio. Finally, evaluate KRIs at a technical level as an integral part of the monitoring process—are they able to predict? PKRI ⇔ LI prediction is one direction, while LI ⇔ PKRI is another direction.

A KRI system is a classic cybernetic system with feed-forward and feedback elements. KRIs help business process owners manage by reducing the number of measures they need for feedback and feed-forward. So, distinguishing between KRIs and RIs and utilizing PKRI ⇔ LI environmental consistency confidence can help to reduce information overload. Herbert A. Simon explains this as follows:

What information consumes is rather obvious: it consumes the attention of its recipients. Hence a wealth of information creates a poverty of attention, and a need to allocate that attention efficiently among the overabundance of information sources that might consume it.[12]

By giving business process owners a clearer focus on key operational risk drivers, they can commission further activities to mitigate them. The PKRI ⇔ LI approach is a dynamic process, not a project to develop a static set of KRIs. This means that a team, possibly aligned with other "scientific" management approaches such as Six Sigma, needs to be constantly cycling through an iterative refinement process over a time period. This implies cyclical methodologies for environmental consistency confidence, such as Z/Yen's Z/EALOUS methodology, illustrated in Exhibit 22.7.

CONCLUSION

Environmental consistency confidence and the PKRI ⇔ LI approach is part of a more scientific approach (hypothesis formulation and testing) to the management of risk in financial institutions.

> *Modern [organization] theory has moved toward the open-system approach. The distinctive qualities of modern organization theory are its conceptual-analytical base, its reliance on empirical research data, and, above all, its synthesizing, integrating nature. These qualities are framed in a philosophy which accepts the premise that the only meaningful way to study organization is as a system.*[13]

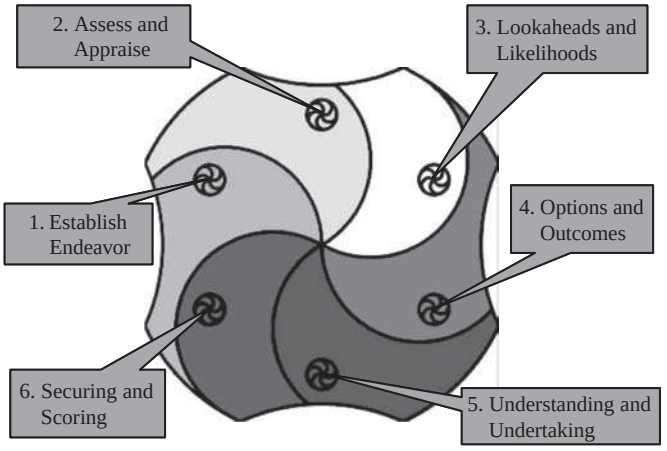


EXHIBIT 22.7 Z/Yen's Z/EALOUS Methodology

At its root, environmental consistency confidence means building a statistical correlation model to predict outcomes and using the predictive capacity both to build confidence that things are under control, and to improve. Good and bad correlations should raise good questions. Today's KRI should be tomorrow's has-been, as managers succeed in making it less of an indicator of losses or incidents by improving the business. Likewise, managers have to create new KRIs and validate them. Regulators should be impressed by an environmental consistency confidence approach, but vastly more important is improving the business and reducing risk by putting statistics and science at the heart of financial risk management.

BIBLIOGRAPHY

Beer, Stafford. (1966). Decision and Control: The Meaning of Operational Research and Management Cybernetics (New York: John Wiley & Sons, [1994 ed.]).

Mainelli, Michael. (2004, May). "Toward a Prime Metric: Operational Risk Measurement and Activity-Based Operational Risk Costing." RMA Journal (special ed.) pp. 34–40, Risk Management Association; www.zyen.com/Knowledge/Articles/toward_a_prime_metric.pdf.

Mainelli, Michael. (2005, June). "Competitive Compliance: Manage and Automate, or Die." Journal of Risk Finance 6(3): 280–284, Emerald Group Publishing Limited; www.zyen.com/Knowledge/Articles/competitive_compliance.htm.

Open Systems Group. (1972; 1981). Systems Behaviour (New York: Harper & Row).

Vapnik, Vladimir N. (1998). Statistical Learning Theory (New York: John Wiley & Sons).

NOTES

1. http://en.wikipedia.org/wiki/Hypothetico-deductive_model.
2. Jonathan Howitt, Michael Mainelli, and Charles Taylor, "Marionettes, or Masters of the Universe? The Human Factor in Operational Risk" *Operational Risk* (a Special Edition of the *RMA Journal*): 52–57, The Risk Management Association (May 2004); www.zyen.com/Knowledge/Articles/marionettes.pdf.

3. Michael Mainelli, "The Consequences of Choice." *European Business Forum* 13 (Spring 2003): 23–26, Community of European Management Schools and Pricewaterhouse-Coopers; www.zyen.com/Knowledge/Articles/Consequences%20of%20Choice%20-%20EBF%2002.03%20v4.1.pdf.
4. Nassim Nicholas Taleb, "The Fourth Quadrant: A Map of the Limits of Statistics." An Edge original essay, *Edge* (September 15, 2008), www.edge.org/3rd_culture/taleb08/taleb08_index.html.
5. Matthew Leitch, e-mail correspondence, 2006.
6. Jules Henri Poincaré,"La Valeur de la Science (1904), from *Value of Science,* translated by G. B. Halsted (Mineola, NY: Dover, 1958).
7. Basel Committee on Banking Supervision, "Sound Practices for the Management and Supervision of Operational Risk." Bank for International Settlements (December 2001).
8. See note 4.
9. Michael Mainelli, "Finance Looking Fine, Looking DAPR: The Importance of Dynamic Anomaly and Pattern Response." *Balance Sheet* 12(5) (October 2004): 56–59, Emerald Group Publishing Limited; www.zyen.com/Knowledge/Articles/looking_dapr.htm.
10. Marcello G. Cruz, *Modeling, Measuring and Hedging Operational Risk* (Hoboken, NJ: John Wiley & Sons, 2002).
11. Michael Mainelli, "Liquidity: Finance in Motion or Evaporation?" Gresham College lecture, London, England, September 5, 2007; www.zyen.com/Activities/Events/Gresham%20College%2013%20-%20Liquidity%20-%20Finance%20in%20Motion%20or%20Evaporation%20v2.1%20-%20for%20email%20v1.0.pdf.
12. Herbert A. Simon, "Designing Organizations for an Information-Rich World." In Martin Greenberger, ed., *Computers, Communication, and the Public Interest* (Baltimore: Johns Hopkins Press, 1971), 40–41.
13. Kast and Rosenzweig in Open Systems Group, 1972, p. 47.

CHAPTER 23

Quality in the Front Office: Reducing Process Variation in Trading Firms

Andrew Kumiega, Ph.D., and Ben Van Vliet

INTRODUCTION

In the new millennium, where just about any trading firm can quickly develop and test thousands of strategies and package successful ones into investment products, business processes are one of the last remaining determinants of competitive advantage. A privileged location on the trading floor doesn't matter in global electronic markets. Since all players have equal access to order flow on electronic exchanges, proprietary risk models, trading algorithms and technologies are quickly reverse engineered. "What's left as a basis for competition is to execute your business with maximum efficiency and effectiveness, and to make the smartest business decisions possible."[1]

Many financial disasters can be clearly explained through the application of standard quality control methodologies. Société Générale lost billions on a rogue trader. A proper quality control system would have produced log files and charts showing the number of trades and change in notional value per day, per person, and per account. Such charts would have uncovered large increases in manual override trades and triggered out of control alerts. Long Term Capital Management's algorithms began deteriorating, producing less and less return with more and more variation. Convinced that the underlying distributions were stable, instead of shutting down the machine they scaled it up, deploying it across markets to generate the higher returns. At All First Bank, a rogue trader manipulated value at risk (VaR) calculations in a spreadsheet. Auditing would have uncovered the fraud.

The goal of controlling quality and producing consistent results is the reduction of process variation. For example, reducing the variation of the output of the trading algorithms. The concept of reduction of variation can be applied to algorithms used for trading as well as to algorithms used to monitor machines that make parts using computer numerical control (CNC). Root cause analysis, in the form of a fishbone diagram, will uncover problems—errors in calculation, lack of benchmarking, versioning, not to mention rogue traders.

The front office needs quality (or its most recent incarnation Six Sigma). The question is how to operationalize quality in so fast-paced an environment. The

answer is through methodologies that control problems that arise during the research, evaluation, knowledge and personnel management, development, and monitoring of quantitatively driven systems and software. Good methodologies provide a consistent framework for making the smartest decisions. Quite simply, Six Sigma techniques can be used to mitigate all forms of risk—operational risk, project risk, credit risk, even market risk.

Most front-office trade selection models are encapsulated in software, be it Excel, Matlab, or C++. Yet some would have us believe that the current financial crisis is not related to information technology (IT). But here is the question: if software encapsulates a bad model, is it then bad software? The answer is yes. A corollary to this question is: can good software testing uncover scientific errors? Again, the answer is yes. Good software development practices mitigate model risk.

The issues in process errors are a result of poor software. Given a model, which is assumed to be correct, say a credit default swap (CDS) pricing model, only rigorous software testing methodologies will uncover the bug. CDS pricing models, though mathematically correct (after all, we can derive them and, furthermore, they have been published in leading academic journals), contain bugs. We now know that these models are horribly incorrect.

Management and regulators should be concerned with the process of assessing a trading or risk management group's ability, or maturity, in creating repeatable processes. This is the fundamental problem with the industry. In our estimation, this is what caused the current financial crisis.

DEVELOPMENT METHODOLOGY FOR QUANTITATIVELY DRIVEN PROJECTS IN FINANCE

We apply quality techniques to front office research and development by way of a 4-stage, 16-step development methodology.[2] Our methodology borrows from the traditional Six Sigma process, define, measure, analyze, improve, control (DMAIC)[3]; and the design for Six Sigma, define, measure, analyze, design, verify (DMADV),[4] to model the process of development of quantitatively driven projects in the front office to ensure the proposed system meets specifications. Our four stages revolve around benchmarking quantitative methods, data cleaning, technology design, and process monitoring respectively. Our framework differs from the Six Sigma and Agile methodologies as due to the heavy research component in front-office development. Unlike standard methodologies, where there are clearly defined methods and goals up front, in the front-office development the methods and goals are fuzzy, or poorly defined, and solutions can be highly complex. Most of these solutions need to be researched and replicated prior to moving forward.

While one may believe that quality does not apply in certain situations, it is important to be sure this decision will lead to better performance. Nevertheless, we do not expect, nor do we advocate, that anyone or any firm follow our methodology exactly. You should modify and adjust our approach to suit your own culture. We do hope you will draw from the concepts presented to design processes that work in your front-office culture (see Exhibit 23.1).

The stages of our methodology include four stages—design, testing, implementation, and monitoring (DTIM)—where the iterations (or spirals), as shown in

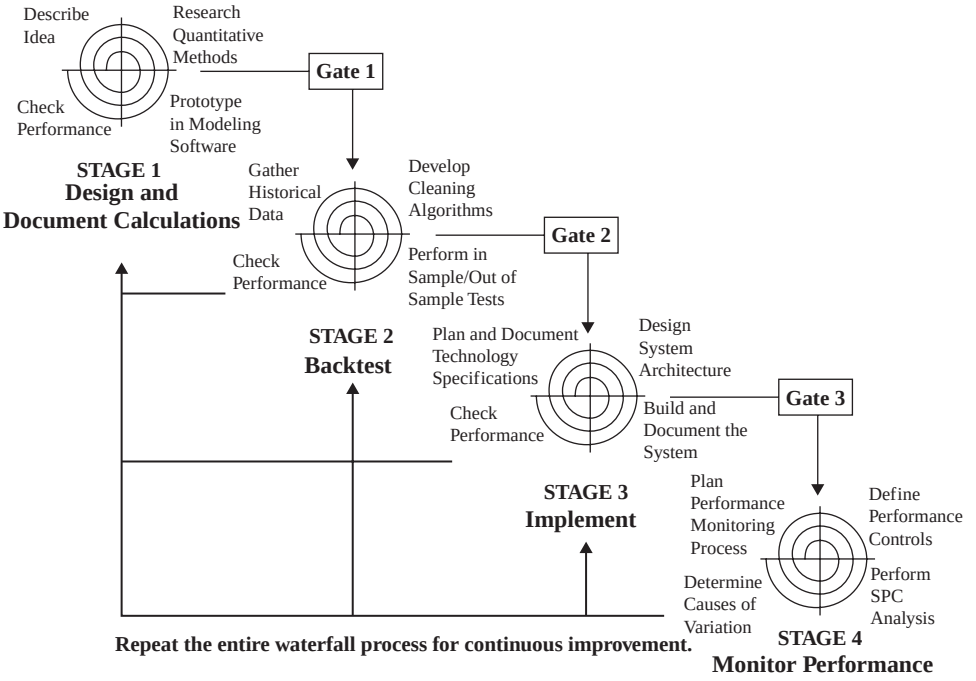


EXHIBIT 23.1 Development Methodology for Quantitatively Driven Projects in Finance

Exhibit 23.2, in each stage devote time to four steps—plan, benchmark, do, check (PBDC):

1. *Plan*. Determine the problem to be solved, gather information, and then plan and document a course of action to solve it.
2. *Benchmark*. Research and compare alternative solutions to arrive at best practices.
3. *Do*. Carry out the best practice course of action.
4. *Check*. Check to see if the desired results were achieved along with what, if anything, went wrong, and document what was learned. If results are not satisfactory, repeat the spiral using knowledge gained.

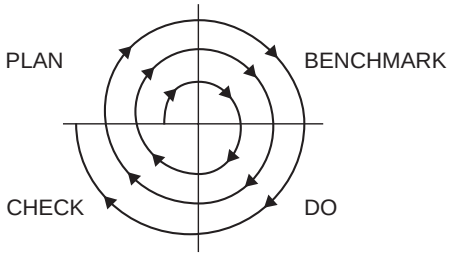


EXHIBIT 23.2 Four-Stage Methodology

Our PBDC framework applies Deming's PDCA (plan, do, check, act) methodology and Motorola's DMAIC methodology for improving business processes. Our model adds benchmarking to these Six Sigma models to emphasize the heavy quantitative research component in finance.

Benchmarking in finance consists of critical comparison of available quantitative methods, data cleaning algorithms, technological implementation, and risk management methods that will yield a competitive advantage. Unlike standard software or manufacturing models, where there are clearly defined methods and goals, in trading system development the methods and goals are fuzzy, or poorly defined, and the solutions will be highly complex. Most of these solutions need to be researched and replicated prior to moving forward. Furthermore, without benchmarking, firms cannot know if methods either derived in-house, or those provided by vendors, are correct.

At the completion of each stage is a gate meeting, where management can check whether or not the business reason for developing the system is still valid. A well-organized gate in the model should make a decision to:

1. *Go.* Go on to the next stage of the waterfall.
2. *Kill.* Kill the project entirely.
3. *Hold.* Hold development at the current stage for reconsideration at a future date.
4. *Return.* Return to a previous stage for additional research or testing.

Essentially, at each successive gate, management must make a progressively stronger commitment to the project. In the end, well-run gate meetings will weed out the losers and permit worthwhile projects to continue.

After completing the fourth and final stage, the methodology requires repetition of the entire four-stage waterfall for continuous improvement, a continuous feedback loop.

Stage 1: Research and Document Calculations

There are two problems with planning in mathematical modeling. First, most financial engineers prefer to start programming, in Excel or Matrix Laboratory (Matlab), immediately instead of creating a software development plan. Second, most project managers lack the quantitative expertise to create a plan for model-driven software, so they rely on the financial engineers to do it. Consequently, many front-office projects never have a formal development plan, a scenario that substantially lowers the probability of success.

Many front-office people realize that the right algorithm for the job may not be the most sophisticated one. Often, the most advanced algorithms cannot be practically implemented in a production environment, or outputs of a model may be too sophisticated for the customer. Research needs a firm business structure to ensure the new models and new software tools solve the customer's needs. Planning research and, furthermore, building and maintaining a proprietary library of unique quantitative methods are keys to the long-term success of any firm in the business of implementing mathematical models to gain a competitive advantage.

Best Practices for Research Top management and financial engineers must determine the design and development inputs relating to the model-driven product requirements, that is, the functional and performance requirements, the performance metrics, the applicable regulatory requirements and laws, and information derived from previous similar designs. These inputs need to be reviewed to assure they are complete and not conflicting.

While most financial engineers, traders, and portfolio managers are conceptual thinkers, trading systems are linear constructs, and planning, documenting, and communicating of model-driven system details must be done linearly. As a result, good researchers plan and conduct their work only with the help of writing. Documenting research forces clarity and understanding with the purpose of communicating specific ideas linearly. Over the course of their activities, successful financial engineers do more than just photocopy source documents, they also write up what they find, keep notes, outlines, summaries, commentary, critiques and questions, maintain a catalog of documentation and sources.

The research step should be a survey of all the relevant mathematical and logical models. Algorithms should be prototyped and benchmarked with sample data. Prototyping allows for clarifying of calculations that will expose inconsistencies in specifications; rapidly evaluating alternative methods; delivery of intermediate, working versions to end users for feedback; clear definition of data requirements; requirements for graphical user interfaces; and development of a working application for regression testing. Research also shows that prototyping leads to improved morale because progress is visible, lower defect rates because of better requirements definition and smoother effort curves, reducing the deadline effect. With respect to prototyping, we also recommend prototyping and testing the riskiest parts first, using team-oriented testing and inspection and documenting prototypes internally.

Front-office personnel are notorious for skimping on quality assurance practices, such as design reviews, walkthroughs, and inspections; some try to make up for lost time by eliminating the testing schedule. A decision to ignore techniques that find defects is tantamount to a conscious decision to postpone corrections until later stages when they are more expensive and time consuming.

Code inspection (and this means Excel, Matlab, and Visual BASIC for Applications [VBA] code) does not entirely eliminate errors; error-reduction methods never eliminate all errors, but a round of team-oriented logic inspection is likely to eliminate 60 percent to 80 percent of errors. That is why we condone iterative development. The more passes over the research stage 1 spiral, the more opportunities to remove errors. We recommend creating a policy requiring comprehensive prototype testing, where such testing should consume 25 percent to 40 percent of prototype development time. We use the term *well-defined* to indicate that a model has been fully prototyped in modeling software and fully inspected and tested. Prior to these steps, a model is merely a dream.

In total, the goal of the research process is to speed the path to the design or application of the best practice algorithms.

Stage 2: Back-Test

Successful algorithm analysis and design necessitates research into past and current markets as a way to analyze and validate the model—a process called *back-testing*.

A back-test is a simulation and statistical analysis of a product's inputs and outputs using historical data.

Prior to building and implementing a new model, developers must test it over a relatively large set of historical data and preferably for a large sample of scenarios. This means building a customized database. While it may seem elementary, investigating the availability of data is very important; because required data may either not exist at all or is prohibitively expensive.

We recommend using Statistical Process Control (SPC) and root cause analysis to identify when and why the underlying data processes have fundamentally shifted. The first step of root cause analysis should be an analysis of economic cycle data around the time the algorithm stopped working.

The focus has to be on making money. At the end of this stage, management must determine (at Gate 2) the potential financial benefit from implementing the project in formalized software and whether or not to proceed.

Best Practices for Data Quality Data quality matters. In some cases, the quality of data may be the determining factor of competitive advantage; it can make or break your front office systems. We categorize data into price, valuation, fundamental (or financial statement), calculated, and economic data. In every type, data, usually purchased from vendors, can range from very clean to very dirty, where there is usually a positive correlation between the price and the quality of data. Using high-quality, more expensive data almost always pays off in the long run, though even high-quality data will not be perfect. When purchasing data, we recommend you evaluate the need and cost of data as well as the experience and reputation of each data vendor.

Best practices for vendor-supplied data include knowing your data and your data vendor; use one consistent source for fundamental data if at all possible; establish data requirements long before making a purchase, and complete a data dictionary and data maps. Data mapping is the process of laying out data elements and conversions between disparate data models, between data sources and destinations.

Benchmarking data-cleaning processes focuses on improving the performance of the front office systems. A best practice for one system, though, may not be a best practice for every system, because each will have its own unique input data. Data problems will affect each system differently. Data problems include bad, or incorrect, data, formatting problems, outliers, and point-in-time data problems.

Whatever the methods used to clean bad data or deal with problems, data cleaning algorithms must be shown to operate on both live-time and historical data. Cleaning algorithms that cannot be performed in real-time prior to model input should not be used on historical data, or else the cleaned, historical data will skew back-testing results. We also recommend that you analyze distributions graphically, clean historical data to correct errors, while maintaining the original, dirty data source in its original form, and normalizing fundamental data across the appropriate grouping. (Normalization reconstructs fundamental data according to identical accounting rules. It is the real service you pay a data vendor for.)

Stage 3: Implement

If prototyping and testing have been completed in previous stages, developing the application in a programming language should be a straightforward march through a

requirements specified by the prototypes. Building a working product from this point is project management. Programming the quantitative algorithms in real code will ensure proper error handling and stability of the system. Furthermore, developers can test-market and check performance of algorithms, data feeds, and usability with beta users.

A product requirements specification (PRS) document should fully define the functionalities and performance requirements of the product. (Notice that the contents of this document have largely been defined over the previous stages.) The PRS document will allow a development team to quickly build the product with the correct features and functionalities and to the proper specifications.

The PRS, along with all software and hardware architecture documents, should be reviewed by all members of the team, including the researchers that designed and built the prototypes. This is often not done, however. Developers often start programming first, and document second. Without a clearly written document and sample code, programmers may make serious mathematical errors when converting Excel or Matlab prototypes into the software code. While these errors should eventually be caught during regression testing against the prototype, they are better prevented through proper, up-front documentation process.

Prerelease acceptance testing will ensure that the software implementation meets the needs of the end users. Prerelease testing should uncover any design flaws prior to full release. The second purpose of this testing is to allow the management time to use the monitoring tools and determine what additional tools need to be built to properly manage the risk embedded in the product.

Stage 4: Monitor Performance

New models and algorithms are new products. These products require constant monitoring. We recommend that periodic reports be generated to show product performance. These reports should present the performance metrics and statistics, and provide documentation regarding the performance of the algorithm relative to its expected performance proved in the back-test. Furthermore, reports should present a determination of the causes of variation from the expected performance and an action plan to deal with those variations.[5]

Processes for monitoring and reporting statistics and risk factors must be implemented. These reports will enable users to know whether or not the product is working within specifications. There are many reasons for deviations in performance at this stage. For example, while performance should be identical to the back-test, error handling and data cleaning in real time may produce unpredictable outcomes. Monitoring should statistically measure the deviations between the algorithm's outputs in the prototype, the back-test, and the working system.

For example, the comparison of quantitative trading strategies to index or peer group benchmarks is well understood. Human traders and money managers and rule-based systems should be judged on their ability to generate excess returns (or lower risk) over and above the benchmark. All competitive endeavors need benchmarks. Without a benchmark, every decision has equal risk-to-reward requirements, which is in direct contrast with Six Sigma principles. The most advanced application of benchmarking a trader's (or trading system's) risk to reward is through attribution analysis, though many other Six Sigma tools exist—Ford 8D, DOE, Fishbone analysis, analysis of variance (ANOVA)—and could be applied to help traders and trading

systems become exceptional. The application of Six Sigma in manufacturing has led to very advanced quality tools to reduce waste and variation through root cause analysis. A good example of this is inventory control and forecasting. The inventory turnover ratio for a product line can be benchmarked against an industry average or an old algorithm, a full understanding of the effectiveness of a new algorithm means analysis across many (not just one) product lines.

Over its life cycle, SPC reports and quantitatively focused tools and products will help front-office personnel understand how their quantitative systems are performing relative to agreed-upon design criteria. After SPC analysis is completed and an understanding of all the sources of variations has been achieved, one final set of tools for process refinement must be built.

WATERFALL PROCESS FOR CONTINUOUS IMPROVEMENT (KAIZEN)

Think of the steps in the methodology as a continuous, never ending spiral of continuous improvement, or kaizen. When applied in the front office, a continuous improvement strategy involves management, traders, financial engineering, programmers, and IT staff (and maybe even marketing and salespeople) working to make small, incremental improvements. It is top-level management's responsibility to cultivate a professional environment that engenders continuous improvement, to focus efforts on eliminating waste. Intelligent leadership should guide and encourage teams to continuously improve profitability, to increase efficiency, and reduce costs. Through small innovations from research and entrepreneurial activity, firms can discover breakthrough ideas.

CONCLUSION

The application of Six Sigma and benchmarking in the front office will result in competitive advantage through maximized efficiency and effectiveness of quantitative research and development, through higher-quality data and software practices. Better business processes will lead to new and better trading strategies and to new products being produced more quickly. Better business processes will lead to better investment decisions. And, finally, better business processes will reduce the probability of a major loss due to operational failure or adverse market movements.

Six Sigma benchmarking can be operationalized in the front office through iterative development for continuous improvement to extend the life cycle of quantitative systems. Firms that embrace Six Sigma will succeed at the expense of those who don't. And those who don't will only in retrospect learn the value of quality.

While we have focused on trading firms for our example, the processes, lessons learned, and best practices are much the same for most all organizations.

NOTES

1. Thomas H. Davenport and Jeanne G. Harris.*Competing on Analytics: The New Science of Winning* (Boston: Harvard Business School Press, 2007), 8–9.

2. Andrew Kumiega and Ben Van Vliet, *Quality Money Management: Process Engineering and Best Practices for Systematic Trading and Investment* (Oxford, UK: Academic Press/Elsevier, 2008).
3. Peter S. Pande, Robert P. Neuman, and Roland R. Cavanagh, *The Six Sigma Way*. (New York: McGraw-Hill, 2000).
4. Kai Yang and Basem S. El-Haik, *Design for Six Sigma* (New York: McGraw-Hill Professional, 2008).
5. W. Edwards Deming, *Out of the Crisis* (Cambride, MA: MIT Press).

CHAPTER 24

The Root Cause of the Global Financial Crisis and Corporate Board Reforms to Prevent Future Failures in Risk Management

Anthony Tarantino, Ph.D.

INTRODUCTION

The world is now suffering through one of most painful economic crises in history. The financial liquidity crisis sparked by the subprime mortgage meltdown may provide the nucleus to reform corporate governance in a way to promote much improved financial risk management, and risk transparency.

BACKGROUND TO THE GLOBAL FINANCIAL CRISIS OF 2007–2009

What is now recognized as a global financial crisis began on February 8, 2007, when HSBC announced the first of many industry write-downs related to subprime mortgages. As an example of the magnitude of the growing scandal, the current write-downs of major banks are 10 times those that occurred with Enron. UBS analysts have projected total losses could reach $600 billion.[1] The *Wall Street Journal* references predictions from some economists of $1 trillion in losses or about seven percent of annual U.S. economic output. More recent loss estimates project international bailouts at over $3.4 trillion, but the final number will continue to grow into 2009. At the $4 trillion level, this would be eight times the loss from the savings and loan crisis between 1986 and 1995, and about four times the size of the Japanese bank crisis two decades ago.[2] The $4 trillion level does not capture the loss in homeowner equity and the human misery from millions of home foreclosures, job losses, and lost investments.[3]

While Enron and related scandals of the early 1990s shook shareholder confidence, hurt millions of investors, and caused the loss of thousands of jobs, it did not cripple international economic growth, force millions of people out of their homes, prevent students from receiving loans, and cause major job losses. The U.S.

savings-and-loan crisis that began in 1986 resulted in $160 billion in losses but had little international impact. The losses from the East Asian financial crisis of 1997 were very significant, but did not occur under more advanced corporate governance and risk management that is now in force in the leading economies. The major scandals in the European Union (EU), Parmalat, and Ahold (2001–2003) had only a minor impact on the overall economies of the EU member states.[4] Finally, the U.S. stock option scandals of 2005–2006 while implicating over 100 U.S. companies, do not appear to have a major domestic or international impact, except to set a very bad example for countries like China that are considering stock options as a share-based compensation approach.

There is now no argument that this is largest financial crisis since the Great Depression of the 1930s. So it is appropriate to look into the real or root causes of the global financial crisis caused by catastrophic failures in financial risk management, and to do so in manner to move beyond regulatory recommendations which are often shortsighted over reactions following major scandals and sometimes end up doing more harm than good, especially in damaging economic growth.

As a Six Sigma black belt with 30 years of operations and compliance experience I am trained to look for the root causes of risk management problems and to recommend permanent corrective actions, which is very much different from the many well-researched published accounts of the tactical and strategic causes of the crisis.

WHY THIS CRISIS DESERVES CLOSE SCRUTINY

The following aspects of this crisis deserve special scrutiny and may provide the means to reform corporate board governance to provide much more robust financial risk management:

- The crisis occurred under the full force of the improved internal control requirements of the U.S. Sarbanes-Oxley Act and within one of the most highly regulated industries: financial services.
- The frequency of U.S. destructive scandals and their corresponding regulatory reactions is increasing. They are causing ever growing grief in spite of claims of very strong corporate governance and risk management. For sure, the United States has the most expensive and complex regulatory structures and highest litigation costs in history but, according to the World Bank, does not enjoy the best corporate governance among leading economies. Others, such as Australia, Canada, and the United Kingdom enjoy higher governance rates, and have avoided waves of destructive scandals, without paying record high audit, underwriting, and litigation costs found in the United States.[5]
- This crisis is causing a great deal of avoidable economic and human misery. Earlier crises hit either real estate or stock markets. The crisis in one sector drove money into the other. This time, both have been hit at the same time.
- The crisis is adding to the loss of U.S. leadership. The United States is no longer the guiding light in human rights or corporate and environmental governance.

Here is a short list of U.S. problems that relate to board governance:

- Marquee scandals with global ramifications.
- The comply-or-go-to-jail approach of Sarbanes-Oxley–like regulations.
- The highest litigation costs in the world.
- Delays in adopting the principle-based International Financial Reporting Standards (IFRS), now in force in most of the leading economies.
- The one year lag behind the EU in adopting the Basel II capital accords for banking.
- The resistance to adopting Solvency II accords for the insurance industry, now in force in the EU.
- A terribly complex tax code which invites gaming.
- The lack of green initiatives, now in force in the EU, such as Restriction of Hazardous Substances (RoHS), Waste from Electric and Electronic Equipment (WEEE), and Registration, Evaluation and Authorisation of Chemicals (REACH) to compel the recycling of chemical, electrical, and electronic waste.
- The lack of EU-like privacy protection initiatives that provide assurances that personal information and e-mail communications are not shared without our permission.
- The rest of world—both friend and foe—blame the United States for the global mess we are in.

THE ROOT CAUSE OF CATASTROPHIC FAILURE IN FINANCIAL RISK MANAGEMENT

The first question is: why America? What is unique in the American experience that has now caused such a major series of crises tied to financial risk management failures over the past 20 years—savings and loan, Enron, stock options, and now subprime? The last two occurred after passing very costly regulatory reforms with very demanding audit standards, resulting in a doubling of audit fees and legal costs. First, let's looks at the usual suspects: greed, stupidity, fraud, and corruption.

Greed

Many in the EU argue that Americans are too greedy, with shareholders demanding much higher growth rates than in the EU. It is also charged that Americans are cowboy capitalists. This is a popular phrase to condemn the very entrepreneurial spirit that has little patience with traditional approaches and models to making money. One could argue that Enron was the case to make this point. Enron's executives were the darlings of Wall Street and the business news media. They were held up as examples to be admired and emulated. More astute observers note that the real cowboy capitalism and scandal existed at Arthur Andersen, their prestigious auditor. A more balanced way of looking at this is that American entrepreneurism and willingness to take on greater risk has resulted in historically higher growth rates than in Japan or Europe, which comes with the higher potential for risk management failures.

Stupidity

The entire scheme behind subprime loans was based on two very dubious assumptions, the first being that it is an acceptable risk to make heavily leveraged loans to individuals with poor credit histories and lower incomes than would have been traditionally approved. In the great majority of cases, loans were made without the traditional 10 percent or 20 percent down payments as a means to hedge or mitigate the risk. Often, credit histories were not even checked. The second assumption was that home prices would continue to rise to sustain the process. While home prices have continued to rise over the past century in the United States, there have always been cyclical swings with downturns in prices. Robert Shiller's analysis shows inflation-adjusted home prices increased 0.4 percent per year from 1890 to 2004, and increased to 0.7 percent per year from 1940 to 2004.[6] This trend was shattered starting in 2000 when prices increased by about 80 percent over the next five years. For the subprime process to continue, rising housing prices and cheap credit both had to continue unabated—something that has never happened continuously in history. The rationalization was that these highly risky mortgages could be bundled up and sold.

There was also plenty of stupidity on the part of buyers of these loans. I recall a recent visit to Las Vegas for a conference in which my cab driver admitted that he and many of his fellow cabbies had each purchased two to four homes on speculation with incomes far below six figures. When I asked him how they were doing on their investments, he said they were typically underwater on their loans and looking at foreclosures no matter how much overtime they worked in order to make their payments. I have examples much closer to home, in my family and among close friends—folks who pride themselves in being sophisticated investors and got burned as badly as the cabbies.

Finally, there was stupidity on the part of U.S. federal and state regulators. With the catastrophic failure of risk management and regulatory oversight, one could argue that the Federal Reserve, Securities and Exchange Commission (SEC), and state regulators all failed to head off obvious lapses in sound banking conduct, stop predatory lending practices, and prevent the abuses in selling mortgaged-backed securities which are now failing at an astounding rate. So the Fed and the SEC failed in their primary charter—to protect investors, maintain fair, orderly, and efficient markets, and facilitate capital formation.

A larger debate is being waged over the Fed than its failure to regulate banks. It is whether its reduction of interest rates to artificially low levels was irresponsible and the ultimate root cause of the subprime crisis. It is argued that this created an irresistible opportunity that overwhelmed risk management, common sense, and common decency.

Fraud

Fraud may be defined as a deliberate deception designed for gain by hurting the interests of another person.[7] Fraud was widespread in the subprime meltdown, with lenders deceiving borrowers, appraisers inflating home value assessments, and borrowers and lenders conspiring to falsify loan applications, credit histories, and bank statements. Fraud may have also existed in selling investments tied to packaging

subprime loans if it is proved that the buyers were mislead to the underlying risks. The litigation process is just kicking into high gear, will take years to play out, and promises to dwarf Enron-era law suits.

Corruption

Corruption may be defined as the abuse of a position of trust for dishonest gain.[8] Corruption is yet to be proven, but a candidate that is bound to be examined is the undue influence lenders made on state and federal lawmakers to prevent the passage of more stringent mortgage controls. The Bernard Madoff scandal, which broke at the end of 2008, exposes the SEC to potential charges of corruption, or at least incompetence, because of Mr. Madoff's connection with the agency and the warnings about his alleged Ponzi scheme received for over eight years.

So we can concede the point that greed, stupidity, and fraud all played a part, but they are not the root cause of the crisis, scandal, and failure of risk management. We can never eliminate them in the human experience, but we can improve political and corporate leadership which will keep these age old sins in check. The recommendations that follow are no guarantee, but would have gone a long way to create a better balance between risk and opportunities, keep shortsighted greed in check, and promote the globalization of markets.

HOW TO PREVENT FUTURE FAILURES IN FINANCIAL RISK MANAGEMENT

There are a number of factors in how corporate boards are currently staffed, structured, operate, and rely on risk frameworks that damage their ability to provide robust financial risk management:

- The current structure of U.S. corporate boards creates what Harold Innis described as a bias of communications in which minority opinions are suppressed. With one dynamic and charismatic individual holding the positions of chief executive officer (CEO) and chairman of the board (CoB), financial risk management tends to take a back seat to the pursuit of opportunities. Separating the two positions has proved successful in the United Kingdom, European Union, and Australia and could provide greater checks and balances between risk and opportunities.
- Risk committees exist at the board level in only a small proportion of financial service firms and are virtually nonexistent in nonfinancial services. A risk committee would give risk management a much better seat at the table of corporate decision making.
- Since about 85 percent to 90 percent of directors are white males, with an average age of 59, their background and perspective does not well represent their major stakeholders—employees, customers, suppliers, and stockholders, especially in global firms.[9] Increasing diversity has proven to improve company performance and should help to broaden risk management perspectives.
- The risk frameworks and related audit standards corporate boards rely on are inadequate to prevent significant breakdowns in financial risk management.

Subprime occurred under the full force of the Sarbanes-Oxley Act (SOX), the Committee of Sponsoring Organizations of the Treadway Commission (COSO) framework, and with many firms meeting the Basel II capital accords. A viable risk framework needs to be quantitative in nature so that risks can be ranked and prioritized allowing boards to focus on the significant few risks that represent the greatest exposure. Such a system can be fairly simple and applicable to organizations of all sizes and complexities. While there has been major progress in improving financial reporting and transparency, this does not translate into improved risk reporting and transparency. Therefore, organizations should use the improved risk framework we described earlier and report on their risk exposure and their plans to mitigate their most significant risks. Improved risk transparency will allow investors, regulators, analysts, and auditors to compare peer organizations against each other and against industry best practices.

- The global crisis has elevated the topic of excessive executive compensation even higher. It was widely debated prior to the crisis but is now a favorite topic far beyond business circles. We argue that excessive pay is a symptom of a bigger problem—the flawed system of executive succession that demands charismatic CEOs to be recruited from a very limited pool of external candidates and then to perform near miracles as organizational saviors. CEOs in this role have been forced to take unprecedented and sometimes unrealistic risks. The charismatic CEO is the product of the shift from managerial to investor capitalism, and from long-term dividend investors, to short-term stock appreciation investors. Reform is therefore no simple matter, but boards can change the succession process to improve longer-term growth and stability.
- Finally, we recommend that organizations institute score cards for the areas we have described here. Simply put, that which is measured tends to improve.

Prevent One Individual from Holding the Positions of CEO and CoB

The first reform is to adopt the U.K. model that prevents one person from holding the position of CoB and CEO. Under this model, the CEO is also not permitted to ascend to the CoB position. Many of the leading economies of the world have embraced this model, and ironically this was the U.S. model prior to World War II. The United Kingdom with its Combined Code/Turnbull Guidance and Australia with its ASX 10 Principles have been leaders in this movement. The Combined Code describes the relative roles and responsibilities of the CoB and CEO as follows:

- There should be a clear division of responsibilities at the head of the company between the running of the board and the executive responsibility for the running of the company's business. *One person cannot hold both positions and establishes the CEO as the head of company operations.*
- No one individual should have unfettered powers of decision. *Decision making is shared between the CoB, the board, and CEO.*
- The chairperson is responsible for leadership of the board, ensuring its effectiveness on all aspects of its role and setting its agenda. *This clearly charges the CoB with assuring the viability and direction of the board.*

- The chairperson is also responsible for ensuring that the directors receive accurate, timely, and clear information. *This helps make the case for more effective communication of the company's risk profile and risk appetite.*
- The chairperson should ensure effective communication with shareholders. This is a critical difference from the United States where the CEO is the chief public face of the company.
- The chairperson should also facilitate the effective contribution of nonexecutive directors in particular and ensure constructive relations between executive and nonexecutive directors. *This has been a major weakness in many companies. A strong audit and risk committee (discussed in the next section) must be independent in order to be effective.*
- The roles of chairperson and chief executive should not be exercised by the same individual. *This is the major difference from the U.S. model.*
- The division of responsibilities between the chairperson and chief executive should be clearly established, set out in writing, and agreed to by the board. *This helps to minimize conflicts, confusion, and power struggles.*
- The chairperson should on appointment meet the independence criteria set out in a later section. This is also a critical difference from the U.S. model and allows the chairman be more responsive to company stakeholders.
- A chief executive should not go on to be chairperson of the same company. This prevents in breeding. The skill sets of the two positions are such very much different and this rule prevents a CEO from scheming for the CoB title. It also reduces the chances of the CoB's undermining the CEO out of fear of their ascension.

The pros and cons of this have been argued for years. Those against the prohibition argue that there is no evidence that company performance improves with two individuals holding the two posts. But advocates argue that there is no evidence that company performance is hurt by the separation.[10,11]

Advocates of combining the roles may cite organizational theory to argue that performance can only be optimized when one person exercises complete, unambiguous, and unchallenged authority. They contend that this provides one public face with a clear company mission. Advocates for splitting the roles may cite principal-agent theory to argue that performance can only be optimized by separating the decision making process with the CEO acting as the decision manager and the CoB as the decision controller.[12]

Performance evaluations of U.S. corporations that switched from a combined to a split model are not always a valid indicator. Many corporations make the change during periods of stress in which the CEO/CoB was replaced for poor performance. In the example of Washington Mutual, the largest U.S. savings and loan only made the split in June 2008 after CoB/CEO Kerry Killinger had lost his credibility and was blamed by the board for destroying the company. Killinger was fired three months later and the 119-year old firm failed in September—the largest bank failure in U.S. history. So an evaluation would show a huge decline in performance after the board removed Killinger as CoB, but the split had no causal impact on WaMu's performance. It is not unusual for U.S. corporate boards to split the ownership of CEO and CoB while forcing out an incumbent CEO, and then to recombine the positions when a new CEO is installed.

EXHIBIT 24.1 FTSE 100 versus S&P 500 Five-Year Percentage Change

One simple way to compare the two models is to look at the relative performance of the major companies listed in the United States (combined model) and United Kingdom (split model). The newest version of the United Kingdom's Combined Code dates to 2003 and mandates a CEO/CoB split. The Financial Times Stock Exchange 100 (FTSE 100) is a stock market index of the 100 most highly capitalized corporations on the London Stock Exchange. The Standard & Poor's (S&P) 500 is a stock market index of 500 large-cap corporations in the United States. Exhibit 24.1 compares the five-year performance history of the two markets. There are no clear advantages for either index and the governance models they represent. In our *Governance, Risk, and Compliance Handbook,* we provide World Bank statistics in which the United Kingdom has consistently scored higher corporate governance and avoided the major scandals that have plagued the United States over the last two decades. Therefore, the British model has historically offered competitive growth rates with the United States, superior corporate governance, and fewer marquee scandals.

The U.S. and Japanese automotive industry presents a head-to-head comparison of the two models. Toyota and Honda, the two top Japanese makers, have pioneered hybrid technology and performed fairly well during the worst economic conditions for carmakers in decades. Both organizations separate the roles of CEO and CoB. (At Toyota, power is shared by a chair, vice chair, and president.) The big three U.S. automakers, GM, Ford, and Daimler/Chrysler, along with the other large Japanese automaker, Nissan, centralize all leadership powers in one individual. Exhibit 24.2 uses Yahoo! Finance to compare the five-year stock performance for the six organizations. While Toyota and Honda stock has barely broken even during the global financial crisis, GM, Ford, Chrysler, and Nissan have lost roughly 60 percent to 90 percent of their stock value. The U.S. big three have performed very poorly during the global financial crisis—GM came close to bankruptcy at the end of 2008, GM and Chrysler have sought government bailouts.

There are some notable exceptions to the combined U.S. model. Of the members of the Dow Jones Industrial 30, six companies operate under a split model: Alcoa, Citigroup, Intel, Microsoft, United Technologies, and Walt Disney. These six represent global leaders in materials, finance, technology, and entertainment.

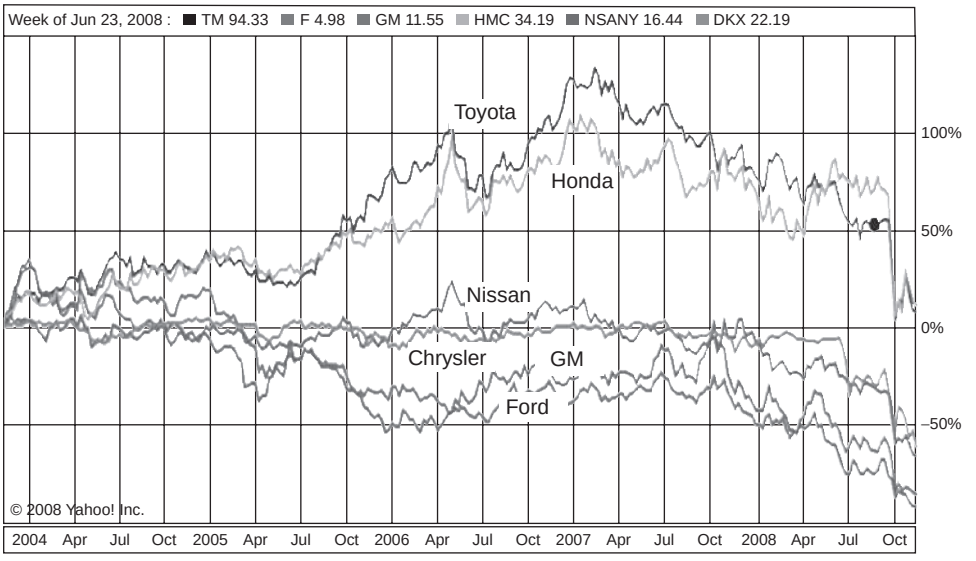


EXHIBIT 24.2 Five-Year Stock Comparison: U.S. and Japanese Automakers

A collateral benefit of splitting the roles between two individuals could be to reduce the large disparity in executive compensation between the United States and the rest of the world, which we discuss in greater detail in a later section. In an ideal environment, compensation committees would be made up of independent directors reporting to a board of directors led by a CoB who is not also the CEO. When a CEO is also CoB, there are obvious pressures on compensation committees. By splitting the responsibility, board-level compensation committees will be less likely to be dominated by one all-powerful person holding both positions.

Separating the two will permit each to focus on critical company objectives—operations and meeting financial targets on the part of the CEO and oversight and the voice of stakeholders on the part of the CoB. It is needed to stimulate much enhanced board governance in which risks and opportunities are more rationally balanced. With an independent CoB, boards can meet and deliberate free of all-powerful and charismatic CEOs who have taken on an imperial presence in the United States.

Create a Risk Committee Reporting to the Board of Directors

The second reform is the creation of a risk committee reporting to the board of directors, as suggested in Australia's ASX 10 Principles of Board Governance. Audit, nominating, and compensation committees are now mandated in many leading nations' company laws. Just as audit committees are typically mandated to be made up only of independent directors and include financial experts, a risk committee should be independent and include risk experts. This would not have guaranteed the prevention of subprime, but would have given a much stronger and independent voice—one not as prone to be sucked into the bias of communication, group think, and shortsighted thinking that punishes opposition.

The role of chief risk officer (CRO) is growing in many companies, led by financial services. While the chief financial officer (CFO) is cited by directors of over 70 percent of company board members as responsible for informing them of risk issues, a growing number of companies are now citing the CRO as the person with primary responsibility—over 16 percent of financial companies, up from virtually zero just a few years ago.[13]

A 2006 survey by the Conference Board indicates wide variations in the quality of risk management from company to company based on feedback from directors who serve on multiple boards, and fewer boards seem to have a well-established risk management process. The survey also found that only 54 percent have clearly defined risk tolerance levels, 47.6 percent of the boards rank key risks, and only 42 percent have formal practices and policies in place to address reputational risk.[14]

According to the survey of Fortune 100 companies, about two thirds of corporate boards place board risk responsibility in the audit committee, but recommends assigning risk management, not associated with financial reporting, to another a separate committee. This committee would then coordinate its efforts with the audit committee, providing improved operational aspects of enterprise risk management. The survey found that risk management is shared with another committee in 23 percent of companies.[15]

Over 15 percent of financial service companies have established separate risk committees at the board level. Outside of financial services, the number drops to less than 4 percent.[16] These percentages will need to increase dramatically to provide the champion to better balance risk with opportunities.

Some general guidelines as to the composition and responsibilities of a risk committee follow:

- It should be made up of at least three members, a majority of which are of nonexecutive directors. This will maintain its independence.
- At least one member should also be a member of the audit committee. This will help to coordinate the two committee's activities.
- At least one person must be a risk expert. As more risk experts become available, it would be advisable to increase this number.
- The chairman of the committee must be a nonexecutive director. This also helps to maintain its independence.
- Overall risk management ownership should reside with corporate boards. They should use best practice frameworks, regulatory requirements, and competitive market forces to guide their risk management decision making.
- The risk committee exists to assist boards in assessing the different types of risk to which the organization is exposed.
- The organization's senior management has the primary responsibility for executing the organization's risk management policy.
- The risk committee should exercise oversight, and must provide evidence about the organization's risk management policy.
- The members of the risk committee need to have direct access to, and receive regular reports from, executive management.
- The risk committee should learn of the actual risks and the control deficiencies in the organization.
- They need to help the board define the risk appetite of the organization.
- They have the duty to exercise oversight over management's responsibilities.

- They review the risk profile of the organization to ensure that risk is not higher than the risk appetite determined by the board.
- They also monitor the effectiveness of risk management functions throughout the organization.
- They ensure that infrastructure; resources and systems are in place for risk management and are adequate to maintain a satisfactory level of risk management discipline.
- They need to periodically monitor the independence of risk management functions throughout the organization.
- They also review the strategies, policies, frameworks, models, and procedures that lead to the identification, measurement, reporting, and mitigation of material risks.
- They review issues raised by the organization's internal audit that impact the risk management framework.
- They ensure that the risk awareness culture is pervasive throughout the organization.
- Finally, they fulfill its statutory, fiduciary, and regulatory responsibilities. This is usually the most difficult task.

Increase Board Diversity

The third reform would be to increase the diversity (female, African-American, Hispanic, and Asian) membership on boards. Many U.S. companies have been moving in this direction for some time seeing it as much more than improved social responsibility and as a means to improve shareholder value by expanding the perspective of corporate boards. The CEO of Sun Oil, Robert Campbell, was quoted over 10 years ago as saying, "Often what a woman or minority person can bring to the board is some perspective a company has not had before—adding some modern-day reality to the deliberation process. Those perspectives are of great value, and often missing from an all-white, male gathering. They can also be inspiration to the company's diverse workforce."[17] The arguments for increased diversity include the following:

- Corporate diversity promotes a better understanding of the marketplace, and its corresponding risks. A more diverse marketplace (suppliers, customers, investors) warrants a more diverse board. This will increase the ability to penetrate these markets and avoid risk land mines.
- Diversity increases creativity, innovation, and more effective problem solving, all of which are key to balancing risks with opportunities. As Robinson and Dechant noted in 1997, "Attitudes, cognitive functioning, and beliefs are not randomly distributed in the population, but tend to vary systematically with demographic variables such as age, race, and gender."
- Diversity enhances the effectiveness of corporate leadership. While board homogeneity promotes quicker consensus, it results in a narrow perspective. A more diverse board will take a broader view resulting in improved decision making, including risk management.[18]

There is new research that helps to explain something that most all parents of teenagers understand—boys take greater and sometimes more foolish risks than girls. Researchers at England's Cambridge University discovered that elevated testosterone

levels in males leads to greater risk taking—sometimes resulting in greater gains and, conversely, to greater losses. They recommended that banks and other financial systems as a whole add more women and older men to their boards and risk management practices.[19] This is not to emasculate management, but to bring a better balance of risk taking and risk sound risk management.

The Conference Board of Canada published a study in May 2002 of the role of women on corporate boards. The study notes a direct correlation between increased female board membership and improved corporate governance. In boards with three or more women, over 90 percent of boards advocated conflict-of-interest guidelines. This compares to less than 60 percent of boards with only male members. In boards with two or more female members, about three quarters of boards conducted formal board performance evaluations. This compares to less than half of boards with only male members. The study also found that boards with increased female membership tend to provide formal board orientation programs and formally limit board authority.[20]

There is evidence that increased diversity improves company performance as measured by the Tobin Quotient (Q = Market value/Asset value) in Fortune 1000 companies. Carter, Simkins, and Simpson conclude their 2003 study: "After controlling for size, industry, and other corporate governance measures, we find statistically significant positive relationships between the presence of women or minorities on the board and firm value, as measured by Tobin's Q." They also found that the proportion of minority and women directors increases with size of the firm size but decreases with increases in higher numbers of inside board members, and that firms committed to increasing the number of minorities on their boards also have more women on their boards and vice versa. Their results provide critical evidence of a positive relation between the value of a firm and the diversity of its corporate board.[21]

In spite of this compelling evidence, female membership on boards continues to lag even in economies in which half of college graduates and postgraduates are women and in economies in which women make up 30 percent to 40 percent of business executives. In the United States, women hold about 11 percent of board positions. With the exception of Scandinavia and some of the emerging East European economies, most European economies lag as well. Norway (currently at 32 percent) and Spain (currently under 5 percent) have imposed quotas to increase female board membership.

Exhibit 24.3 is a chart by the EU Commission that shows the large gap between the percent of female executives and female board members.

The EU Commission survey does provide a potential target for organizations looking to increase female board membership. It is the percentage of female executives. Using this criterion, female board membership would roughly triple over its current rate. In the United States, it would include people of color in executive positions. It could also be argued that it makes sense to expand the survey to include those in executive positions in labor, government, and higher education.

Improve the Risk Framework

While the COSO framework has brought a much-needed framework to financial reporting and transparency, it needs to be supplemented to include a stronger risk framework.

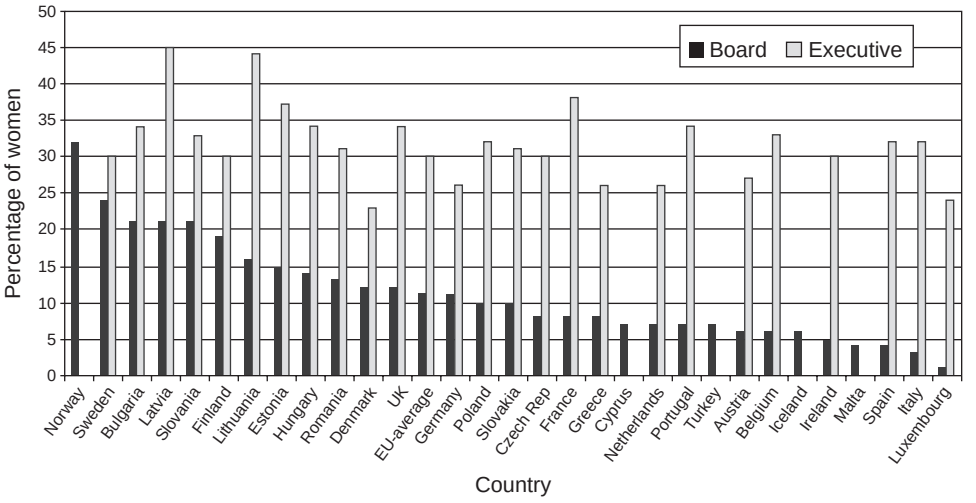


EXHIBIT 24.3 EU Commission, Percent of Women in Executive Positions and as Members of the Boards, 2006

Weaknesses in the COSO Framework An indirect but important cause of the global financial crisis and other breakdowns in risk management is the risk limitations of the COSO framework that the Sarbanes-Oxley Act (SOX) and many other corporate governance protocols are based upon. COSO created an integrated internal control and risk framework in 1992 that was updated to include enterprise risk management (ERM) in 2004. The 1992 framework identified five interrelated components:

1. Control environment
2. Risk assessment
3. Control activities
4. Information and communication
5. Monitoring

Besides the COSO framework, American auditors utilize Statement of Accounting Standards (SAS) 31, "Evidential Matter," which created five general classifications of assertions. The job of auditors is to look for the following:

1. Existence
2. Completeness
3. Valuation
4. Rights and obligations
5. Presentation and disclosure[22]

The categories and classifications of COSO and SAS 31 are not the problem, but the lack of a means to prioritize the audit process is a major problem. The U.S. audit reforms that call for a top-down approach will still fall short, because there is

no viable means to apply a numerically weighted framework that would permit it to score, rationalize, and prioritize risk as to its:

- Financial impact
- Likelihood of occurrence
- Ability to be detected

As a result, many insignificant risks tend to require the same degree of scrutiny as the few major risks that can derail or destroy a company. To put this in perspective, consider that many larger organizations are required to audit several hundred internal controls. In global organizations, there are typically over 1,000 controls. But there are no major shortcuts for internal controls that have a minor (nonmaterial) impact on financial reporting.

The criticism of COSO is not new and includes the Institute of Management Accountants (IMA), which charges that COSO was never designed to fully meet the mandates set by the SEC and U.S. SOX. The brutal truth is that COSO is the creation of the audit industry and not used by risk managers as their weapon of choice.[23] Another problem is COSO's organizational status. It is not governed by a national or international regulatory body. It is a committee, with no legal status, and lacks funding to support research, training, and education.[24]

COSO is considered the safe choice because of its widespread adoption and the framework of choice by countries such as China adopting their own versions of Sarbanes-Oxley. China's new Standard for Enterprise Internal Controls makes it very clear that it is based on the COSO framework.

A summary of the arguments for the need to supplement COSO with an improved risk framework include:

- *COSO did not prevent the global financial crisis.* U.S. SOX is the most extensive use of the COSO framework, with the highest internal and external audit costs. Even after four years of use under Section 404, COSO did not identify the huge risks that organizations faced. The European country laws, with the exception of France, are based on the COSO framework, and also failed to identify the risks.
- *COSO does little to prevent most major fraud and risk failures.* Most fraud occurs at the executive and board level—above the internal controls COSO is designed to address. Enron, the marquee scandal of the last decade, had nothing to do with failures in internal controls, but with the abuse of off-balance arrangements. The lack of transparency to huge risk exposures in the financial industry occurred under audited and certified financial results under the COSO framework.
- *COSO's financial disclosure provides poor risk disclosure.* Some financial organizations with the highest marks for their financial disclosure under U.S. and EU corporate laws and have collapsed or been severely damaged due to their failures in risk management. In the most notorious examples (e.g., Leman Brothers), they failed with no warnings to investors or regulators. Under section 409 of the Sarbanes-Oxley Act, organizations are required to declare material weaknesses within a very few days of discovering such pending disasters. Clearly, the process has failed to provide risk exposure warnings.

A Viable Supplement to the COSO Framework—Risk Quantification and Scoring There is a means to supplement COSO that will provide a much improved framework over internal controls and by extension improved risk management. Using even a simple system of risk quantification and rationalization along with improved risk management oversight would have at least given subprime mortgages, mortgage-backed securities, and credit default swaps a great deal more exposure at the board, management, and auditor level. It may not have prevented the crisis, but could have reduced its impact.

Using the three criteria mentioned above, such a system of risk quantification and ranking could work like this:

- All risks, both internal and external, are ranked by three criteria: financial severity, likelihood of occurring, and ability to detect (other criteria may be added or substituted to fit an organization's environment).
- Assign a numerical value (e.g., 1 to 10) to the three criteria for each risk.
- Add the three criteria together.
- List the risks in their descending risk score.
- Focus the greatest attention on those items with the highest risk scores.

In the large majority of environments, Pareto's 80/20 rule will apply in which less than 20 percent of items (those with the highest risk scores) will represent over 80 percent of the risks an organization faces. Historically, accountants informally applied a five percent rule in which balance sheet items that represented less than five percent of total value were not an area of focus.

Such a commonsense approach using our risk scoring will allow organizations to focus on the very significant few items that represent the great majority of risks they face. Such a system would benefit from establishing and publishing industry-specific risk frameworks. In any case, organizations' own ranking should be subject to review by auditors, regulators, and rating agencies.

In order to be viable, this system would need to be incorporated at the management and board levels and would supplement the COSO framework. With the movement toward the IFRS as the global accounting standard, there is a need for a much-improved global auditing and risk framework. Ideally, a new risk framework would incorporate Six Sigma. Once an organization has prioritized its risk items, Six Sigma black belts would be ideal to lead the projects to attack the most dangerous risks. Their proven problem-solving and project management techniques will be invaluable in the process.

Summary In summary, for every process, there is typically some associated risk that requires an internal control. For processes that impact financial reporting, internal controls are subject to financial audits that evaluate their effectiveness. The COSO framework is heavily auditor biased and needs to be supplemented with a risk-based framework created and facilitated by risk experts. Auditors have a critical role in establishing the rules for the audit and conducting audits that will restore the confidence of investors and other stakeholders, but a new framework and disclosure process that evaluates and exposes the most significant risks an organization faces is essential.

Provide Risk Transparency Reporting

All publicly held companies must periodically report their financial results. Financial statements consist of four elements: a balance sheet, income statement, statement of retained earnings, and statement of cash flow. Together, they provide a comprehensive snapshot of the short-term and long-term financial position of a company, but do little to provide transparency to the short-term and long-term risk exposure. During the global financial crisis, major financial services companies failed after submitting financial results attesting to their financial well-being. This occurred under the most rigorous U.S. and EU reporting requirements, with many of the EU firms also following increased capital and internal control requirements of the Basel II capital accords.

While financial reporting is extremely complex, risk reporting can be very simple. It would include a descending list of the highest risk exposure to an organization with a rationalization for the assessment and the mitigants to the risk that are in place and/or planned. The Basel Committee has established a viable hierarchical categorization for operational risk. In order to compare risk self-assessments from one organization to another, it would be helpful to apply the Basel categories and subcategories. With the coming of extensible business reporting language (XBRL), and using the Basel categories, it will be possible to easily compare peer organizations within industry sectors.

Comparing the risk assessments will at least provide insights into the risk thinking of an organization. It will be valuable to compare peer organizations and look for similarities and differences in their assessments. Weaker organizations can benchmark their risk assessments against the industry leaders. But, history provides warnings that industry consistency in risk assessments is no guarantee of success. In the 1960s, the three big U.S. automobile makers believed their primary risks came from their U.S. competitors. The real threat came from Japan, with its superior quality and manufacturing efficiencies. This only became clear to them decades later.

Reform Executive Succession and Compensation

Background Rakesh Khurana, a Harvard Business School professor, in his 2002 book *Searching for a Corporate Savior: The Irrational Quest for Charismatic CEOs,* describes the U.S. change from owner-based, to managerial, to investor-driven capitalism that has occurred over the last 100 years and fundamentally chained the risk appetite of corporate America.[25] In the early twentieth century, business owners were compelled to delegate control to professional managers as they sold a growing portion of their companies to shareholders and investors to finance their continuing growth. Managerial capitalism proved very successful with its highly trained and experienced managers until the early 1970s, when corporate profits and U.S. competitiveness declined.

Historically, investors had little control over corporations in which they invested. In the mid-1980s, investors—especially large institutional investors—became more vocal in their demands on boards and executives to improve corporate performance. As a result of the increased pressure, U.S. CEOs were three times more likely to be dismissed after 1990 than before 1980.[26]

Corporate directors came to believe that they could exert greater control over external CEO candidates than internal candidates, and viewed external candidates

as a means to satisfy investors, analysts, and the business media. This could be best accomplished by hiring a marquee name as a charismatic savior of the organization.

The rise of investor-based capitalism and frustration with the lackluster performance of incumbent corporate management laid the foundation for what has come to be known as a charismatic or imperial CEO. The traditional organizational man was replaced with a celebrity who demanded celebrity levels of compensation.

Executive salaries soared in this market because boards, investors, analysts, and the business media all mistakenly believed that such a great leader could cure any and all corporate woes. This had two negative consequences beyond higher executive salaries. First, the new CEO was under inordinate pressure to perform miracles. This led to their taking on extraordinary risks, which sometimes resulted in major losses up to and including the demise of the organization. Second, this undermined the need to develop strong subordinate executives who could succeed the CEO and therefore would strive to improve corporate performance.

Executive compensation increases have been dramatic. In 1965, CEOs and CFOs were paid 20 times more than the average worker. The gap in 2007 is now over 300 times and averages $10.5 million for CEOs in the S&P 500.[27,28] The gap in the United States is much larger than in the rest of the world, with U.S. executives making twice as much as their German, French, and British counterparts and four times as much as their Korean and Japanese counterparts.[29]

A basic philosophy in business management is succession planning. Many organizations require incumbents to identify and train their potential replacements. Such measures improve performance and help assure continuity when incumbents leave their positions. By relying on external candidates only, boards have undermined the performance of their own management and raised executive compensation to levels unacceptable to virtually everyone except the executives receiving it.

The level of executive compensation is the most criticized element of this problem, but it is the nature of the compensation that presents the greatest risks to an organization. Before the 1980s, most executive compensation was primarily fixed and in cash. The culture of charismatic CEOs flipped this ratio so that variable is now the large majority of executive compensation and usually share-based.

The share-based nature of the variable compensation is an issue because it is often based on increases in the company's share prices either through stock options or restricted stock, which creates major incentives for executives to take extraordinary measures to jack up share prices. This can lead executives to make short-term measures at the expense of the long-term growth of the organization. In the worst situations, a temporary price increase is generated by manipulation and accounting games in order for executives to exercise options.

The global financial crisis has created very heated public and official outcries against excess executive compensation, especially multimillion-dollar severance packages given to failed executives who led their firms to catastrophic losses. As we noted earlier, executive compensation is a symptom of the charismatic CEO culture, which has resulted in much greater risk taking. Here are some recommendations to reform executive succession and compensation.

Recruit Chief Executives from Within the Organization Reform needs to start with boards accepting that one person, no matter how famous a personality, is not a substitute for a strong management team. A strong management team must be

composed of at least some members who are capable of ascending to the CEO and CFO positions. This change in philosophy will have the benefit of creating greater incentives for senior managers to excel to prove their viability for promotion.

Internal candidates can be much more thoroughly vetted than external candidates, who are often selected through an imprecise and hurried process based on anectdotal information.

The argument that only a charismatic external candidate can fix the major issues an organization faces is a simplistic and emotional response to very complex problems that require the efforts of several key executives, senior managers, and supervisors to solve. Without strong internal candidates, organizations may suffer from lower energy levels, initiative, and innovation.

There were valid issues of inept and caretaker management that plagued the United States in the 1970s and 1980s and led boards to look outside the organization for salvation. For the most part, these issues have been resolved by the demands of the global economy and more demanding institutional investors. If they have not been resolved, boards have failed in their mission.

While hiring an external charismatic CEO may result in a boost in the company stock, this will tend not to last without fundamental improvements. A better investment is for corporate boards to upgrade the senior executive staff to prevent the types of crises that compel boards to go outside the organization for its leadership.

Change the Nature of Executive Compensation As mentioned earlier, traditionally executives received the bulk of their compensation in cash, with a smaller portion coming in bonuses. This has changed in the past 20 years, with more and more compensation tied to share-based compensation. Executives should be rewarded for performance, with the majority of their compensation in cash and a minority tied to longer-term incentives. This can bring more stability to organizations and reduce excessive risk taking without sacrificing long-term growth. Pressure from analysts, the business media, and proactive institutional investors will tend to keep executives very focused and motivated to perform. Boards can always remove executives who fail to live up to expectations.

Stock options are not a viable option in most cases in that they are often tied to short-term incentives. The argument that options are the best means to align the interests of shareholders and executives is flawed and reflects the day-trader mentality of many investors and analysts. Executives have many vehicles to artificially jack up stock prices to maximize their option rewards. These activities may be detrimental to the long-term well-being of the organization.

The United Kingdom has been a leader in the movement away from stock options and other share-based compensation to long-term incentive plans (LTIPs). LTIPs are a reward system designed to improve the long-term performance of executives and employees by providing rewards that may not be tied to the organization's share price. Like stock options, clever executives can and have manipulated LTIPs to work in their favor. Trevor Buck, Alistair Bruce, Brian G. M. Main, and Henry Udueni describe the LTIP manipulation practices in the United Kingdom: "While increasing average total rewards, the presence of LTIPs is actually associated with reductions in the sensitivity of executives' total rewards to shareholder return." They argue that this raises doubts as to their effectiveness.[30]

The best defense against manipulation may be to tie compensation to metrics that are measured and averaged over three or more years and to use accepted best

practices in LTIPs, which we describe in the next section. This helps avoid practices that artificially inflate share prices and ultimately undermine the long-term well-being of the organization.

Apply Best Practices in Executive Succession and Compensation Matsumura and Shin, two professors at the University of Wisconsin–Madison, provide a list of six best practices that should be applicable to any organization seeking to improve its executive compensation practices. We eliminated one, which calls for CEOs to increase their equity in the firm, and replaced it with the recommendation against share-based compensation. We also add one requesting accounting standard bodies to create best practices for LTIPs.[31]

1. **Executive compensation needs to be aligned with the long-term interests of shareholders and with corporate goals and strategies**. *Long-term* is the critical term here to avoid the types of dramatic actions to artificially boost stocks, only to see them decline again when the poor risk management of such actions is realized. As such, executives need to be measured to performance-based metrics that tie to long-term shareholder value, which is balanced against the potential risks.
2. **An independent compensation committee needs to determine the compensation of the top executives.** *Independent* means it is composed of independent directors only. As we argued earlier, this will work best when the CEO is not also the CoB. This prevents the obvious pressure that would fall on even independent directors.
3. **Compensation committees need to thoroughly understand the total costs of the compensation packages they are considering**. This requires accounting support to project the total costs of retirements, severances, travel, and various long-term benefits. For many executives, these costs can run into the millions of dollars. The poor performance of U.S. compensation committees is now common knowledge. In the past, many of them naively believed that stock options were virtually free. Under revised international accounting rules (IFRS), options are now expensed and can have a major impact on company earnings while diluting the value of company shares. This was demonstrated when many U.S. firms had to restate earnings as a result of the stock option back-dating scandal of the past five years.
4. Compensation committees need the services of nonbiased, independent, and experienced advisers to guide them in selecting and modifying compensation packages. Some compensation committees have foolishly relied on external consultants who were retained at the behest of incoming CEOs to justify inflated salaries. Typically, they would point to other inflated executive compensation packages for externally recruited and charismatic CEOs to justify their recommendations. Hopefully, the global financial crisis will make compensation committees more leery of taking such actions, but recruiting internal candidates and preventing CEOs from ascending to the CoB may be the best means to break this cycle.
5. **Companies need to provide complete compensation transparency**. The United States and many EU nations now require more disclosure as to executive compensation. Unfortunately, the disclosure does not always provide transparency to the true costs of a wide variety of benefits and perks. Regardless of the regulations, shareholders deserve full disclosure in an understandable format of the

compensation of the top executives. The failure of the current U.S. regulations can be seen in the huge public outcries over the severance packages given to terminated executives of the major financial service organizations. The disclosure rules did not provide significant insights into the costs of golden parachutes that ran up to $100 million.

6. **Accounting standard bodies need to publish guidelines as to LTIPs.** This will help to eliminate manipulation by executives, allow compensation committees to avoid the mistakes of the past, and facilitate tax and financial reporting. When used prudently, LTIPs may be the best means to align shareholder interests with incentives to company executives. Selecting from a list of approved LTIPs should help to validate the process.
7. **Companies need to avoid stock options and other share-based compensation plans.** In the *Governance, Risk, and Compliance Handbook*, we dedicated a chapter to the dangers of stock options and argue that there are better means to reward executives. Even if all the abuses around back-dating and hiding expenses are resolved, it is still a bad idea that measures employees to a metric over which they have little control. Executives, who can influence share prices, face too many temptations to manipulate events to maximize their option exercise price levels. The IFRS requirement to expense options will end one major abuse, but does not change their inherent problems.

Create and Publish a Corporate Governance Scorecard

There is truth in the old adage "that which is measured improves." We have listed areas in which risk management can be improved. A scorecard will provide an easy means to measure an organization's progress in improving its corporate governance around risk management. In our *Governance, Risk, and Compliance Handbook*, we call for a voluntary approach to SOX section 404, which covers internal controls that impact financial reporting. This includes a scorecard for those that opt in to the program. Organizations would be given a grade based on number of material weaknesses and financial restatements they receive.

A similar program can work for risk management. Most of these proposed reform areas can be given a simple pass/fail grade. Historically, investors and other stakeholders have relied on rating agencies for such indices, but the process has many flaws, which are now becoming abundantly clear—financial institutions failed after receiving very high ratings.

Most of the recommendations are easily monitored and graded. The risk management framework would require an organization to list its descending list of high-risk items and its programs to mitigate these risks. Even if these reforms are embraced, it will be years before they become statutorily mandated in whole or in part. Therefore, a scorecard for publicly listed organizations will be essential to provide the marketplace with the visibility it needs to make more rational investment decisions.

CONCLUSION

The global financial crisis can provide very painful lessons learned to move America forward and the potential for the best of all worlds—fewer and less severe scandals, higher growth, and greater stability.

Root cause analysis typically comes with recommendations for permanent corrective actions. The permanent corrective actions we make here are very attainable with improved government and corporate leadership. Most of our recommendations have been proven within the United States or elsewhere—by America's major trading partners.

The alternatives are very unattractive. Doing nothing virtually assures we will continue to suffer wave after wave of increasingly destructive scandals and crises. This will make the United States and other laggards less likely to attract global capital as other regions enjoy higher growth, improved corporate governance, and fewer marquee scandals. Creating additional but tactical regulations as occurred during previous scandals will invoke the specter Einstein used to define stupidity: doing the same thing over and over again and expecting a different result. In this case, targeted regulatory action will help end abuses behind subprime, but could create other negative consequences, and do little to prevent the next crisis.

It will take a holistic approach with systemic reforms, such as the ones recommended here, to break the cycle we have fallen into—boom to bust to scandal to overreactions in regulations and litigation. At the end of the day, capital will flow to markets that best balance growth with creditability and accountability. These reforms will never completely break the age-old and vicious cycle in which periods of laisles-faire activity with inadequate oversight leads to scandals, and scandals in turn lead to regulatory action. Unfortunately, the pendulum tends to swing too far in each direction—under regulation permitting scandals and crises to flourish to overregulation which stifles growth.

With much higher growth rates in emerging economies and the relative stability and security of the EU, the United States can no longer afford these wide swings between under regulation and overregulation. For the United States to remain competitive in global markets, its goal should be to mitigate these destructive cycles in such a way that reforms are less reactionary and less burdensome, especially to entrepreneurship; in such a way that improved corporate governance better balances opportunities with risk and common decency; and in such a way to prevent the human and economic misery that comes with major crises.

NOTES

1. Abigail Moses and Yalman Onaran, "Financial Firms Face $600 Billion of Losses, UBS Says." Bloomberg.com, February 29, 2008; www.bloomberg.com/apps/news?pid=20601085&sid=anDZQ703DEn4&refer=europe.
2. Carrick Mollenkamp and Mark Whitehouse, "Banks Fear a Deepening of Turmoil." *Wall Street Journal* (March 17, 2008): 1, 12.
3. Robert Winnett, "Effort to Halt Financial Crisis Costs Governments Two Trillion Pounds." Telegraph.com.uk, October 15, 2008; www.telegraph.co.uk/news/3198470/Effort-to-halt-financial-crisis-costs-governments-two-trillion-pounds.html.
4. Anthony Tarantino, *Governance, Risk, and Compliance Handbook* (Hoboken, NJ: John Wiley & Sons, 2008): 13–15.
5. Ibid., p. 919.
6. Wikipedia, "The United States Housing Bubble." http://en.wikipedia.org/wiki/United_States_housing_bubble.
7. Hriskikesh D. Vinod, "Fraud and Corruption," in Tarantino, 2008, p. 121.

8. Ibid., p. 121.
9. David A. Carter, Betty J. Simkins, and Gary W. Simpson, "Corporate Governance, Board Diversity, and Firm Value." *Financial Review* (February 1, 2003).
10. Jay Dahya, "One Man, Two Hats—What's All the Commotion." City University of New York, CUNY Baruch College, Zicklin School of Business, August 2005; http://papers.ssrn.com/sol3/papers.cfm?abstract_id=853006.
11. Maria Carapeto, Meziane A. Lasfer, and Katerina Machera, "Does Duality Destroy Value?" Cass Business School, City University, London, January 12, 2005; http://papers.ssrn.com/sol3/papers.cfm?abstract_id=686707.
12. Ibid.
13. Kay Brancato, Matteo Tonello, and Ellen Hexter, "The Role of the U.S. Corporate Board of Directors in Enterprise Risk Management." The Conference Board, Report No. 1390, June 6, 2006.
14. Ibid.
15. Ibid.
16. Ibid.
17. See note 9.
18. Ibid.
19. Randolph Schmid, "Male Hormone Linked to Irrational Risk Taking." *San Francisco Chronicle,* April 15, 2008, p. D2.
20. Judy B. Rosener, "Women on Corporate Boards Make Good Business Sense." Womens Media.com, May 2003; www.womensmedia.com/new/Rosener-corporate-board-women.shtml.
21. See note 9.
22. See Anthony Tarantino, *The Managers Guide to Compliance* (Hoboken, NJ: John Wiley & Sons, 2006), 147–152.
23. Tim Leech, "COSO—Is It Fit for Purpose?" In Anthony Tarantino, 2008, p. 75.
24. For a detailed evaluation of the shortcomings in the COSO framework, see Tim Leech, 2008, pp. 65–75.
25. Rakesh Khurana, *Searching for a Corporate Savior: The Irrational Quest for Charismatic CEOs* (Princeton, NJ: Princeton University Press, 2002).
26. Ibid., pp. 59–60.
27. Albert R Hunt, "Letter From Washington: As U.S. rich-poor gap grows, so does public outcry," *Bloomberg News,* February 18, 2007.
28. Heather Landy, "Behind the Big Paydays." *Washington Post,* November 15, 2008.
29. See note 27.
30. Trevor Buck, Alistair Bruce, Brian G. M. Main, and Henry Udueni, "Long Term Incentive Plans, Executive Pay and UK Company Performance," *Journal of Management Studies,* 40(7), September 26, 2003, pp. 1709–1727, www3.interscience.wiley.com/journal/118870450/abstract?CRETRY=1&SRETRY=0.
31. Ella Mae Matsumura and Jae Yong Shin, "Corporate Governance Reform and CEO Compensation: Intended and Unintended Consequences." Department of Accounting and Information Systems, School of Business University of Wisconsin–Madison, January 31, 2005.

Index