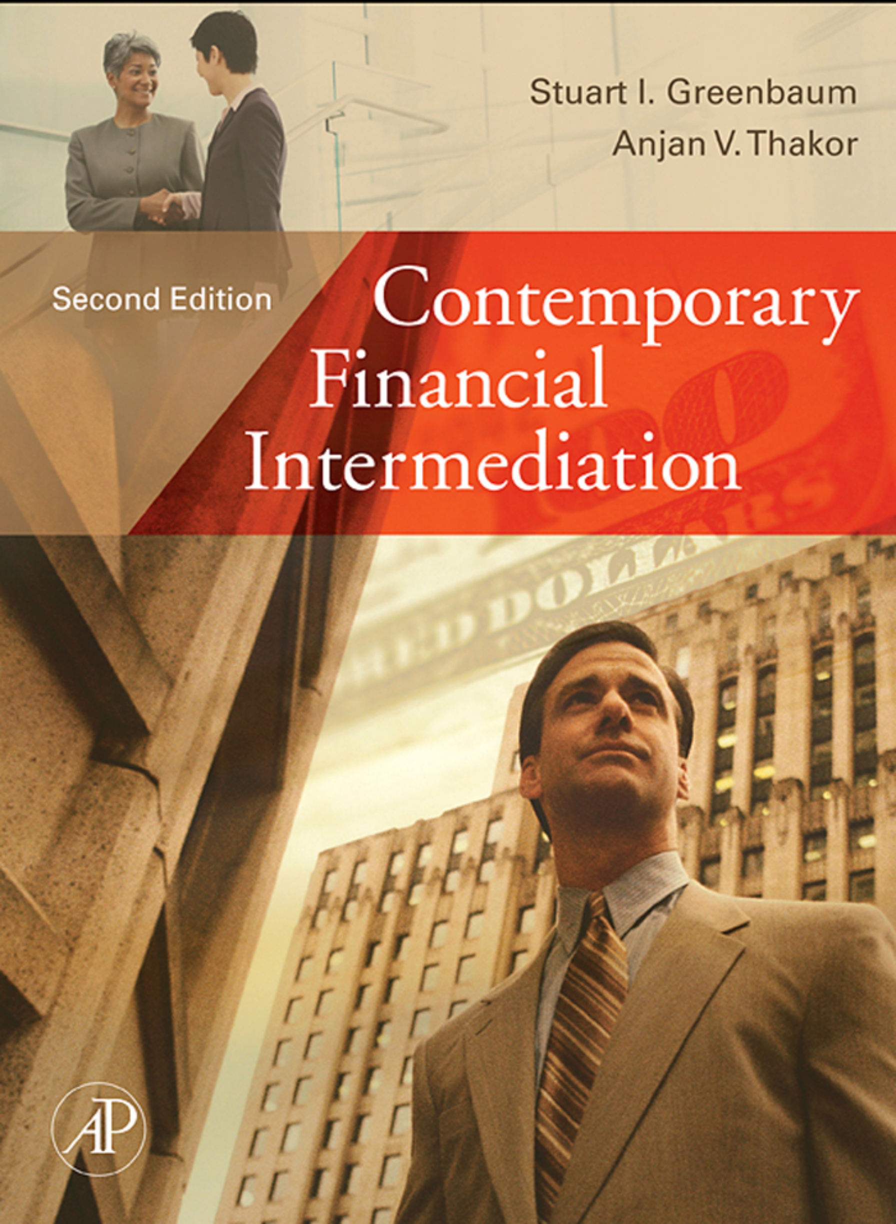
ACADEMIC PRESS ADVANCED FINANCE SERIES

Stuart I. Greenbaum
Anjan V. Thakor

Second Edition

Contemporary Financial Intermediation



Contemporary Financial Intermediation

This page intentionally left blank

Contemporary Financial Intermediation

Second Edition

Stuart I. Greenbaum

Washington University in St. Louis

Anjan V. Thakor

Washington University in St. Louis

ELSEVIER

AMSTERDAM • BOSTON • HEIDELBERG • LONDON
NEW YORK • OXFORD • PARIS • SAN DIEGO
SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO

Academic Press is an imprint of Elsevier

Academic Press is an imprint of Elsevier
30 Corporate Drive, Suite 400, Burlington, MA 01803, USA
525 B Street, Suite 1900, San Diego, California 92101-4495, USA
84 Theobald's Road, London WCIX 8RR, UK

This book is printed on acid-free paper. ♾

Library of Congress Cataloging-in-Publication Data
Greenbaum, Stuart I.
Contemporary financial intermediation / Stuart I. Greenbaum, Anjan V. Thakor. – 2nd ed.
p. cm.
Originally published: Fort Worth : Dryden Press, 1995.
Includes bibliographical references and index.
1. Banks and banking–United States. 2. Financial services industry–United States. 3. Intermediation (Finance). 4. Bank management. I. Thakor, Anjan V. II. Title.
HG2491.G727 2007
332.10973–dc22
2007000154

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

ISBN 13: 978-0-12-299053-3
ISBN 10: 0-12-299053-6

For information on all Academic Press publications
visit our Web site at www.books.elsevier.com

Printed in the United States of America
07 08 09 10 9 8 7 6 5 4 3 2 1

To Elaine, Regina, and Nate
My spiritual lenders of last resort
Stuart I. Greenbaum

To my parents, Lata and Viru
For everything that made this possible
Anjan V. Thakor

This page intentionally left blank

Abbreviated Contents

Extended Contents

Preface

In writing this book we set out to modernize the teaching of bank management at universities and collegiate schools of business. Our goal is to expand the scope of the typical bank management course by (1) covering a broader, but still selective, variety of financial institutions, and (2) explaining the *why* of intermediation, as opposed to simply *describing* institutions, regulations, and market phenomena. Our approach is unapologetically analytical, and we have tried to make analysis an appealing feature of this book. We will consider the book a success if it leads students to not only discover the endless subtlety and plasticity of financial institutions and credit market practices, but also develop an appreciation for *why* these institutions, market practices, and governmental regulations are encountered. The unifying theme is that informational considerations are at the heart of what most banks do.

The novelty of our approach lies in both the analytical orientation and our choice and sequencing of topics. We begin with the questions of why financial intermediaries exist and what they do. We believe that understanding the why of financial intermediation will prepare the readers for the inescapable volatility of the future. Regulations, institutions, and claims will change, but the functional foundations on which financial intermediaries are built will remain basically the same.

Pedagogy

Each chapter (except "A Friendly Conversation" and Chapter 1) begins with a glossary of terms that students will encounter while reading that chapter and will revisit throughout the book. Key nonbanking concepts are discussed in Chapter 1 to provide students with a clear basis on which to proceed. Within each subsequent chapter, we provide numerical examples, laying out each step from idea to solution. Each chapter ends with review questions, and many chapters include case studies to help students appreciate the power of the concepts as well as the complexities.

Moreover, because some chapters contain basic as well as more technical materials, more advanced discussions are isolated in boxes. Interesting, but inessential, information is likewise presented in isolated passages. This provides the instructor with enhanced flexibility in customizing the course.

Organization

The book contains 16 chapters and "A Friendly Conversation." In Part I, the introductory chapter consists of dialogues among three friends about banking in both 1991 and 2007. It is a mix of sound ideas and naiveté. Much of what is discussed in this chapter will be unfamiliar to a student without previous exposure to the subject. The chapter challenges students' knowledge of the issues, and it could be covered the first day of class to obtain students' viewpoints on various issues. We refer back to this conversation throughout the book in end-of-chapter review questions that test the students' expanding knowledge. "A Friendly Conversation" could also be discussed at the end of the course to gauge the changes in the students' viewpoints.

Chapter 1 discusses the key concepts of information economics, game theory, market completeness, options, and other topics we use throughout the book. We recommend that these concepts, which are central to the issues encountered in subsequent chapters, be discussed when needed in the context of subsequent chapters, rather than being dealt with at the outset of the course.

Remaining chapters address eight distinct topics. In Part II, Chapters 2 and 3 examine the functions of financial intermediaries. Chapter 2 describes the variety of financial intermediation and the basic services provided by financial intermediaries. Chapter 3 sets forth the information-based theory of financial intermediation and explains how banks evolved from goldsmiths.

Part III addresses the three basic business risks of banks: interest rate, liquidity, and credit risks. Chapter 4 discusses how these risks are related. Interest rate risk is explained from the vantage point of the arbitrage-free term structure of interest rates (under both certainty and uncertainty). In addition, we consider the importance of information as a source of liquidity risk. Chapter 5 focuses on credit risk and the lending decision. Credit rationing and other lending anomalies are examined in Chapter 6. New in this edition, Chapter 7 covers a few special topics in credit, including syndicated loans, loan sales, and project finance.

Part IV deals with "off-balance sheet" banking. Chapter 8 discusses commercial bank contingent claims, including loan commitments, letters of credit and bankers' acceptances, interest rate swaps, and related contracts like caps, collars, and swaptions. Chapter 9 addresses securitization.

Part V covers the liability side of the bank's balance sheet. Chapter 10 explains particular aspects of the demand deposit contract and also examines deposit insurance.

Bank regulation is covered in Part VI by Chapters 11 and 12. First, we consider the different regulations to which banks are subject, and the economic/political rationale for each. The history of U.S. banking regulation as well as the institutional structure of regulation are examined. Then in Chapter 12, we turn to an analysis of proposals for regulatory reform. In particular, we discuss the 1991 FDIC Improvement Act, and the Basle II Capital Accord adopted in 2004.

In Part VII, Chapter 13 pulls together the key management questions found in previous chapters. It discusses both the day-to-day and the strategic management of opportunities and the three key business risks—interest rate, liquidity, and credit risks—in banking. It also discusses crisis management.

Part VIII deals with corporate control and governance in banking. Chapter 14 discusses bank mergers and acquisitions. In Chapter 15, we discuss issues concerning investment banking.

Finally, in Part IX's Chapter 16, we look to the future, conjecturing about the evolution of banking in the United States and elsewhere. There are three main themes in our discussion: the continuation of globalization in banking, risk management by banks, and international capital regulation.

We believe it will be difficult to cover the entire book in one academic quarter or even one semester. Students for whom this book is intended are not accustomed to thinking about asymmetric information and agency issues, so it takes time to become familiar with the basic concepts. We recommend that the instructor select a subset of topics, keeping in mind that it would probably require two semesters to comfortably complete the entire book. Possible course outlines are included in the Instructor's Manual.

Whatever the approach chosen by the instructor, we hope that this book provides an accessible, if intellectually challenging, rendering of contemporary banking thought. Our own experience in teaching these materials has been rewarding. We hope the same is true for others.

Supplementary Materials

Instructor's Manual/Test Bank/Transparency Master: Initially prepared by Daniel Indro of Kent State University and revised for this edition by Jian Cai of Washington University in St. Louis, the Instructor's Manual includes lecture notes and outlines for each chapter, as well as answers to the end-of-chapter questions and case studies. To offer instructors more flexibility, the Instructor's Manual provides citations of recent articles that instructors can include in their class. Summaries and discussion questions are provided to help incorporate these articles for class discussion. The Test Bank offers approximately 500 questions and problems for use on exams, homework assignments, and quizzes. A set of overheads is also available for all chapters except the first one so that instructors can use these in their classroom presentation.

This page intentionally left blank

Acknowledgments

The second edition of *Contemporary Financial Intermediation* has benefited from the advice of many colleagues and friends. In particular, we would like to thank Neal Stoughton, Chris Hatina, and Terry Wirtel at Washington University in St. Louis for all their hard work in typing the manuscript. In addition, the following reviewers gave of their time and insight to improve the manuscript in its many stages.

Nasser Arshadi, University of Missouri at St. Louis
Sheldon D. Balbirer, University of North Carolina at Greensboro
Jess Beltz, Hong Kong University of Science and Technology
Arnoud Boot, University of Amsterdam
Yuk-shee Chan, Hong Kong University of Science and Technology
Elizabeth Cooperman, University of Baltimore
Robert Eisenbeis, University of North Carolina at Chapel Hill
David Ely, San Diego State University
Mark J. Flannery, University of Florida
Gary Gorton, University of Pennsylvania
John H. Hand, Auburn University
Jim Johannes, University of Wisconsin at Madison
George Kanatas, University of South Florida
G.D. Koppenhaver, Iowa State University
William Kracaw, Pennsylvania State University
Morgan J. Lynge, Jr., University of Illinois at Urbana-Champaign
Loretta Mester, Federal Reserve Bank of Philadelphia
Todd Milbourn, Washington University in St. Louis
Helena Mullins, University of British Columbia
Joseph A. Newman, Northern Illinois University
Ehud I. Ronn, University of Texas at Austin

William Sartoris, Indiana University
James Seward, Dartmouth College
Case M. Sprenkle, University of Illinois at Urbana-Champaign
Ali Tarhouni, University of Washington
Greg Udell, New York University
James Verbrugge, University of Georgia

About the Authors

Stuart I. Greenbaum

Stuart Greenbaum is the former Dean and professor emeritus at the John M. Olin School of Business at Washington University in St. Louis. He is the 2006 recipient of the Lifetime Achievement Award of the Financial Intermediation Research Society. He was named the Bank of America Professor of Managerial Leadership in 2000. Before joining the Olin School in 1995, Greenbaum served for 20 years on the faculty of the Kellogg Graduate School of Management at Northwestern University where he was the Director of the Banking Research Center and the Norman Strunk Distinguished Professor of Financial Institutions. From 1988–1992, he served as Kellogg's Associate Dean for Academic Affairs. Before Northwestern, Greenbaum served as Chairman of the Economics Department at the University of Kentucky, and on the staffs of the Comptroller of the Currency and the Federal Reserve.

Greenbaum has served on fifteen corporate boards. He also served on the Dean's Advisory Council of the Graduate Management Admission Council, on the board of AACSB International—The Association to Advance Collegiate Schools of Business, Executive Committee of the World Agricultural Forum, and on the board of the St. Louis Children's Hospital. He was thrice appointed to the Federal Savings and Loan Advisory Council, and was twice officially commended for extraordinary public service. Greenbaum has consulted for the Ewing Marion Kauffman Foundation, the Council of Higher Education of Israel, the American Bankers Association, the Bank Administration Institute, the Comptroller of the Currency, the Federal Reserve System, and the Federal Home Loan Bank System, among others. He has on numerous occasions testified before Congressional committees, as well as other legislative bodies.

Greenbaum has published two books and more than 75 articles in academic journals and other professional media. He is founding editor of the *Journal of Financial Intermediation* and has served on the editorial boards of ten other academic journals.

Anjan V. Thakor

Anjan Thakor is John E. Simon Professor of Finance and Senior Associate Dean of Programs, Olin School of Business, Washington University in St. Louis. Prior to joining the Olin School, Thakor was The Edward J. Frey Professor of Banking and Finance at the Ross School of Business, University of Michigan, where he also served as chairman of the Finance area. He has served on the faculties of Indiana University, Northwestern University, and UCLA. He has consulted with many companies and organizations, including Whirlpool Corporation, Allision Engine Co., Citigroup, RR Donnelley, Dana Corporation, Anheuser-Busch, Zenith Corporation, Lincoln National Corporation, J.P. Morgan, Landscape Structures, Inc., CIGNA, Borg-Warner Automative, Waxman Industries, Reuters, The Limited, Ryder Integrated Logistics, AT&T, CH2M Hill, Takata Corporation, Tyson Foods, Spartech, and the U.S. Department of Justice. Among many other honors, Dr. Thakor is the winner of the Reid MBA Teaching Excellence Award, Olin School of Business, 2005, and received the Outstanding Teacher in Doctoral Program award for the University of Michigan Business School, April 2003. He has published over 100 papers in leading academic journals in Finance and Economics and books of readings. Besides this book, he has published five other books.

PART ♦ I

The Background

This page intentionally left blank

A Friendly Conversation

Introduction

Before investing in a book, you should ask whether it's really worth the effort. The answer depends on what you bring to the undertaking, To assist you in forming a preliminary judgment, we present a mythical conversation among three reasonably well-informed friends. The conversation concerns banking. It covers a few of the topics that we deal with in this book, but certainly not all of them. To us, this conversation raises more questions than it answers, rather than illuminating any specific issues. Its principal objective is to provide a test of how much students know about banking at the outset, and then perhaps to see how much they have learned in the course. So we recommend that this chapter be discussed in the first week of class to learn students' views, and then perhaps again at the end of the course. We believe that it is difficult to formulate intelligent answers to the questions that are implicit in this conversation without understanding the issues examined in later chapters. But you be the judge.

The Conversation: 1991

The three friends are Alex Appleton, Beth Butterworth and Mike, the moderator. The time is early 1991 and the three friends are engaged in an animated debate about the recently publicized financial crises in the savings and loan (S&L) and banking industries.

Moderator: So, what do you people think? Will we ever really understand what happened to the American banking industry well enough to know what should be done?

Appleton: Well, I think banks and S&Ls were simply victims of the environment. We had an inverted yield curve—long rates were lower than short rates—for a while and this made it difficult for financial institutions to reap their normal profits from asset transformation; you know, I've never believed in the expectations hypothesis. It's a theoretical nicety with no practical relevance. Of course, the increased interest rate volatility didn't help. As if this weren't enough, there was an enormous increase in competition, both domestic and international. These institutions must have felt like they were being squeezed by a powerful vise.

Moderator: And let's not forget those myopic politicians who encouraged banks to take on significant LDC (loans to developing countries) exposure. Do you know how much bank capital was wiped out as a result of LDC writeoffs? It sure puts the European banks at a competitive advantage. Also, all of the deregulation and reducing capital requirements didn't help either. By the way, Alex, I'll give you another reason not to like the expectations hypothesis—it's also wrong.

Appleton: I didn't know that. Are you sure? In any case, it's good to know you agree with me, Mike. But frankly, I'm surprised. Knowing how you and Beth feel about this, I thought I'd get more of an argument.

Moderator: Well, cheer up, Alex. My agreement with you is only partial. I agree that depository financial institutions faced a tough environment during the last 15 years or so. But I also think they could have *managed* their risks more intelligently. For example, they could have reduced the duration gaps in their asset and liability portfolios and made use of contemporary immunization techniques to hedge their interest rate risks. Like some of the investment banking houses, they could have been more innovative in brokerage activities, so that the resulting fee income would have made banks less dependent on the riskier asset transformation activities. Just look at the profits earned by some investment bankers who stripped Treasuries and sold zeros (pure discount bonds) like CATS (Certificates of Accrual on treasury Securities) and TIGRS (Treasury Investment Growth Receipts). No, Alex! The real story runs much deeper than your "passive victims of the environment" explanation. I think banks and S&Ls *exploited* the system and ripped off taxpayers.

Appleton: Mike, you're paranoid.

Moderator: Am I really? More than 50 percent of the S&L failures involved management fraud.

Butterworth: It's kind of amusing to listen to both of you, because neither of you is completely right. Mike, even though fraud was detected in more than 50 percent of failed S&Ls, I believe that the dollar losses due to fraud added up to less than 5 percent of the total dollar losses. So the fraud issue is a bit of a smokescreen. I think the *real problem* is that we designed a banking system in the 1930s and it's outdated.

Moderator: I don't see where you're disagreeing with me, Beth. After all, isn't it tautological to say that a system that allows itself to be exploited by depository institutions is outdated?

Butterworth: Not quite! My point is *not* that the system *allowed* itself to be exploited. Rather, the system *encouraged* depository institutions to do the things that they did. By and large, I don't believe that banks and S&Ls did many things that were not in the interests of their shareholders. Rather than being the victim of exploitation by banks and S&Ls, the system provided the *incentives* for these institutions to engage in the activities you have termed "exploitation." There's a difference between crying foul because a thief breaks into your house while you're away and crying foul after you have invited the thief into your house to carry away your possessions.

Moderator: We may be getting bogged down in semantics here. Could you be more specific, Beth?

Butterworth: Well, I'm referring to the distorted risk-taking and capital accumulation incentives provided by our system of governmental regulation. Risk-insensitive deposit insurance pricing gave endowed banks and S&Ls low-cost put options and created a monstrous moral hazard problem. Regulatory uncertainties artificially pushed up the cost of bank capital and, combined with declining charter values, really exacerbated the moral hazard problem. What we ended up with was a system totally lacking in any sort of incentive compatibility.

Appleton: Beth, most of what you are saying is totally incomprehensible to me. Didn't you tell me the other day that you thought that implementing a risk-sensitive deposit insurance pricing scheme could be a real nightmare? So, why pick on the risk insensitivity of deposit insurance pricing as the culprit?

Butterworth: I still strongly believe what I said then, Alex. But that doesn't contradict what I'm saying now. It's kind of tricky to explain this, but . . .

Moderator: Excuse me, Beth, but I have to leave in a little while, so perhaps we can move on and talk about what can be done to *improve* the system. I read recently that the Treasury Department proposed to reform our banking system. It looked to me like there were some good ideas in that proposal. What do you think?

Butterworth: Well, Mike, it is an interesting proposal, but not everything in it is new. I like the part about regulatory consolidation and dismantling of the McFadden Act restrictions on nationwide branching. I'm not crazy about the elimination of Glass-Steagall—the Banking Act of 1933 that separated commercial and investment banking—because I think it continues to serve a constructive purpose.

Appleton: Frankly, I don't think that the regulatory consolidation proposal goes far enough. I like Henry Gonzales' proposal to consolidate all banking regulation in just two agencies a lot better on that score. I have idea why you are so enamored of Glass-Steagall, Beth. Doesn't it make sense to level the playing field for banks and their competitors?

Butterworth: Alex, it's not that I like Glass-Steagall *per se*. It's just that if you're going to eliminate it, you have to be careful about *how* you do it, and *what* you replace it with. That is where I think the Treasury proposal is lacking.

Moderator: But I thought that the proposal was careful to recommend a hierarchy of capital levels so that only the relatively well-capitalized banks could engage in many of the activities proscribed by Glass-Steagall.

Butterworth: I know, but that's a long way from achieving what I'd like to see. I could explain, Mike, but you have to run.

Moderator (with a wry grin): I appreciate that, Beth. Talking about capital, you know I haven't quite thought through the ramifications of the Treasury proposal in light of the BIS (Bank for International Settlements) capital guidelines that will become effective in 1992.

Appleton: That's simple, Mike. The BIS stipulations are *minimum* levels, whereas the Treasury proposal gives banks choices above the BIS minimum. What bothers me about the BIS guidelines, though, is that they also require banks to hold capital against *off-balance sheet items*. When these items get *on* the balance sheet, there is another capital requirement against them, so aren't we in a sense double counting?

Butterworth: Not really, because there isn't simultaneity involved. I think that with a trillion dollars in outstanding loan commitments alone, the issue of the contingent liability exposure of American banks is something that we just have to come to grips with. The way that RAP (Regulatory Accounting Principles) and GAAP (Generally Accepted Accounting Principles) accounting have dealt with these contingent liabilities has been deplorable. I strongly believe depository institutions should be made to recognize these liabilities *on* their balance sheets, not merely in footnotes.

Appleton: Beth, I think you're getting a bit carried away. Nobody has any idea how these contingent liabilities should be valued, so how do you quantify your exposure?

Butterworth: Speak for yourself, Alex. There *are* valuation models available, although I'll admit they are far from perfect. But even imperfect information is better than none.

Appleton: I think you're wrong on this. What you're saying is closely related to calls for mark-to-market accounting, that the SEC (Securities and Exchange Commission) seems so taken with. I heard recently that the Chairman of the Board of Governors of the Federal Reserve System, Alan Greenspan, sent a letter to the SEC voicing his objection to compelling banks to mark all their assets to market. I think that some pretty knowledgeable people are beginning to recognize the difficulties with market value accounting.

Moderator: Hold it there people. Remember, I can't be here forever. I thought we were discussing banking reform and deposit insurance. Does all this talk about off-balance sheet activities have anything to do with deposit insurance?

Butterworth: That's a good question, Mike. I honestly don't know, but my guess is that contingent liabilities represent a hidden liability for the deposit insurance fund. The more contingent liabilities the banks have, the more risk there is in the banking system.

Appleton: As both of you know, I believe that off-balance sheet activities are the future of banking, so Beth's views on this trouble me. Perhaps she has some evidence to support her claim?

Butterworth: No, Alex, I don't. But I'll research the matter.

Moderator: Well then, I guess it's time to get back to deposit insurance. You know, I went to a seminar the other day and heard someone say that the simplest solution to the deposit insurance problem was *not* to have any federal deposit insurance at all. I couldn't hang around long enough to find out why he said that, but does that make any sense to you?

Butterworth: No. Federal deposit insurance prevents bank runs. If you don't have deposit insurance, then it's possible to have panic runs on banks even though the economy is healthy and has not received any adverse shocks. These can do serious harm to the economy.

Appleton: Well, for once I'm familiar with the theoretical basis for your argument, Beth. But surely, you're not suggesting that deposit insurance is the *only* way to prevent bank runs. What about suspension of convertibility or 100 percent reserve requirements?

Butterworth: Suspension of convertibility won't work as well as deposit insurance in solving the bank runs problem, although as you know, it's been tried in the past. The 100 percent reserve requirements solution will work but I don't like it at all because *fractional reserve banking* is the historical foundation of depository financial intermediation. Something very central is sacrificed in separating the payments and credit-creation functions of banks.

Moderator: In general, I don't like using reserve requirements as an instrument to facilitate bank liquidity. Even the Fed has officially dropped that pretense. But I must admit that I'm at a loss. We seem to be saying here that there is no hope for sensible reform. Is that true?

Appleton: Mike, I don't think I ever intended to say that. One proposal that I am intrigued by is that we eliminate deposits and let bank liabilities reprice like mutual fund shares. Some say this will totally eliminate bank runs, but I'm not so sure.

Butterworth: Nor am I.

Appleton: Another proposal I like is the "narrow bank" concept. We could have federally insured deposits but require that these be invested only in very safe assets, like T-bills.

Moderator: Fine, but as long as you have fractional reserve banking, you're never going to eliminate the possibility of withdrawal risk altogether.

Appleton: That's why you have a lender of last resort, Mike.

Moderator: OK! That's one for you, Alex. But I don't understand one thing. What happens to all of the assets that banks currently fund?

Appleton: No big deal. These can be shifted to the capital market or funded with uninsured deposits.

Moderator: But is such disintermediation or reintermediation necessarily a good thing?

Appleton: I don't see why not. Banks are already securitizing many of their assets, from credit card receivables to mortgages. What I'm suggesting is only a natural extension of that process.

Butterworth: Sure, but there are natural limits to securitization. Besides, even with securitization, the bank acts as an originator. What you're proposing, Alex, is based, I think, on the premise that there is really nothing special about banks.

Appleton: Absolutely! I believe that when you cut through all the bull, the essential role of banks is to act as "lot breakers" and provide simple transaction services. I can't write checks against a T-bill, so I need a bank.

Butterworth: Alex, I couldn't disagree more. Everything that I've read suggests that banks *are* special. Your proposal would destroy a key ingredient of the process by which society allocates capital from savers to investors.

Moderator: It looks to me like we have a *fundamental* disagreement: Why do we have banks and what do they really do?

Appleton: What's to disagree? Ask anybody and they'll say that banks are there to borrow and lend money.

Moderator: That's obvious, but it hardly settles the issue, does it, Alex? After all, borrowing and lending are *not* services. The question is: What are *outcomes* of banks' production of financial services. The question is: What are these less transparent financial services that banks and other financial intermediaries produce? You say that the services are purely transactional, while Beth claims they are much more.

Butterworth: That's a neat way to put it, Mike. You know, Alex, I'm not saying that transactional services are unimportant. My point is simply that the private information and moral hazard problems that banks resolve are also important.

Moderator: And the key is to recognize that we can't sit here and evaluate *how* we should reform banks without understanding the economic function of banks. Only after we understand what it is that banks do, can we ask how alternative reform proposals *affect* the efficiency with which these services will be produced in the future. If the only service of banks is purely transactional, then Alex's proposal makes perfect sense to me. But if Beth is right, then I'm not sure.

Butterworth: Actually, Alex's proposal isn't bad, it's just incomplete. What we should do is have his narrow bank embedded within a larger bank that has the ability to invest in virtually any asset it wants as long as these assets are financed with liabilities that are not federally insured. That way we could have safety without undermining the financial intermediation process. Of course, we would need effective "Chinese Walls" around the narrow bank.

Appleton: Aren't you being a little naïve, Beth? I think that the idea that these Chinese Walls will be impenetrable is naïve. And if they're not foolproof, we're back to square one.

Butterworth: Oh, come on Alex! You're on this "give me perfection or give me nothing" trip again.

Moderator: Alright, let's change the subject. I'm a little surprised that we haven't yet talked about the role of the regulators themselves in all of this. A lot of people have recently been criticizing regulators and accountants as being largely responsible for the S&L mess.

Butterworth: And with good reason too! Regulation not only provided depository institutions with perverted incentives, but the regulators' behavior aggravated the resulting problems. You know, S&L regulators concede now that they knew back in the early 1980s that most S&Ls had negative economic net worth, but chose not to liquidate them.

Moderator: How were they kept alive?

Butterworth: Like zombies, for the most part. RAP and GAAP are wonderfully elastic standards.

Appleton: Well, the problem is that the whole process is so thoroughly politicized. Have you looked at the way that the RTC (Resolution Trust Corporation) is set up? The whole purpose seems to be to mislead the taxpayer about the total cost of the salvage rather than to complete the salvage at the lowest cost.

Butterworth: That's an interesting viewpoint, Alex. I hadn't realized that.

Appleton: That's the first time all evening that Beth has agreed with me. But seriously, I think a big problem is that banks don't really know what to expect from regulators. Uncertainty in regulation is not a diversifiable risk. So it increases the cost of capital for banks and encourages them to pursue activities that require less capital support. I think securitization, swaps, and the variety of off-balance sheet activities that we have observed are all due to this.

Moderator: That's the second outstanding point you have made in a row, Alex. But if banks increase their off-balance sheet activities, might that not lead to less *on-balance* sheet activities? Do you think all this is going to reduce the involvement of banks in lending to business—the traditional role of commercial bank lending?

Appleton: Perhaps, but not necessarily. And what if it does? But I better stop here because my views on this will probably provoke a strong objection from Beth.

Butterworth: I'll let that pass because I want to address your question, Mike. You know over 70 percent of business loans are secured, and collateral has some really beneficial incentive effects from the bank's standpoint. Moreover, it permits the bank to engage in creative loan-contract design, which helps to resolve some thorny informational problems, It also leads to improved bank monitoring of borrowers, which is a key function associated with both secured and unsecured lending. To make a really long story short, I think that business lending is a key component of banks' activities. If regulation discourages this, then I think we'll have seriously weakened the financial intermediation process.

Moderator: If the role of banks in business lending were to diminish, what sort of losses to society do you foresee, Beth?

Butterworth: That's my favorite topic, Mike, so we could be here all night if I get going. But just briefly, I think that banks have developed considerable expertise in originating these loans, designing loan contracts, structuring covenants, including the crafting of collateral requirements, monitoring, and the restructuring of loans for borrowers in financial distress. It would be a shame if the financial system evolved in such a way that these skills would need to be relearned by others.

Appleton: If banks don't do it, someone else will.

Butterworth: I'm sure that's true, but the question is one of comparative advantage and deadweight losses, that is, reinventing the wheel. For instance, take the example of DIP (Debtor-in-Possession) financing. There's nothing in the law that says only banks can provide it, but banks are the biggest players in that market. It's not a mere coincidence.

Moderator: I guess it's not surprising that the DIP financing market has grown so much, given the debt binge of American corporations in the last decade. I personally find the whole debt restructuring process, and particularly the role of banks in it, quite fascinating. But I do find it ironic that banks are engaged in this at a time when borrowers are complaining about credit rationing by banks.

Appleton: I think this concern with credit rationing is overdone. First of all, I don't really believe banks ration credit, and if they did, it would be irrational. I'm not in the habit of worrying about why someone may want to smoke a \$5 bill! Moreover, a borrower who is rational could always go elsewhere. But honestly, I have yet to see a convincing study that shows that banks ration credit.

Moderator: Come now, Alex! Do we need a convincing empirical study substantiating every little truth?

Butterworth: Please don't answer that, Alex. The fact of the matter is that it is possible to explain credit rationing as a rational practice. And this view that a rationed borrower can go "somewhere else" is not surprising coming from you Alex, since you don't believe banks are special anyway.

Moderator: To change the subject, do either of you have any opinion on how American banks are going to stack up against foreign banks in the future?

Appleton: Well, I believe that the Japanese banks are going to be less of a competitive threat than the Europeans.

Butterworth: Why?

Appleton: Because the Americans and the Europeans are better positioned right now in terms of their capital levels. The name of the game is going to be capital. I think the European banks will grow worldwide at the expense of the Japanese and possibly the Americans.

Moderator: And that takes us back to regulatory reform in terms of the potential effects it could have on the future of U.S. banks.

Butterworth: That means we better make sure we understand what it is that our banks do and what it is that we *want* them to do.

Moderator: Amen!

Appleton: Since I agree, this is a good time to say goodnight.

Moderator: Good night!

Butterworth: Good night!

Follow-Up to the Conversation: 2007

The three friends return in late 2007 and strike up another friendly conversation. The topic now is capital regulation and deposit insurance reform.

Moderator: A lot has happened since we last met. The S&L failures are a distant memory, the Glass-Steagall Act is gone, we have had a great deal of experience with the Basel I Capital Accord, the Basel II Capital Accord was adopted last year, the FDIC has plenty of reserves, banks have to keep capital against off-balance sheet contingent liabilities, and so on.

Appleton: That's right. I guess we were not right about everything after all.

Butterworth: I'm sure we're all a bit selective in our memory recall. But let me jump in here and ask you about the replacement of the Basel I with the Basel II Accord.

Appleton: I think it's a good thing. Basel I was a great improvement over what we had before. But Basel II is so much more sophisticated.

Moderator: Which specific aspects of Basel II are you referring to, Alex?

Appleton: The three pillars, Mike. The whole objective of Basel II is to adequately control banking system risk. I like the fact that instead of relying on a single instrument—capital requirements—we will now use three classes of instruments: capital requirements, regulatory monitoring and market discipline.

Butterworth: Yeah, I like it conceptually too, but I'm not sure I'm a believer in it yet.

Moderator: Why not, Beth? Is it just because Mike is so enthusiastic about it?

Butterworth: No, no. I like simplicity, and Basel II is anything but simple. I don't think you can implement all of its features at a large bank unless you have a Ph.D. in finance, statistics, or math.

Appleton: Oh! Come on, Beth. Large banks can hire a quant jock to deal with all the statistical stuff. It's not that big a deal.

Butterworth: Perhaps. But I really liked the simplicity of Basel I. It wasn't perfect and it could be gained by the banks. But overall, I think it worked.

Moderator: OK, folks. I don't think we are going to settle this debate here. I know lots of good arguments on both sides, but let me turn to the merger between Citicorp and Travelers that created Citigroup. I haven't had a chance to ask you about it since this occurred some years ago.

Butterworth: Well, I thought it was a gutsy move to go ahead with the merger since it involved a merger of banking and insurance companies, which was against the law at that time because the Glass-Steagall Act was still in effect.

Moderator: Yes, I know. Although I don't know anyone who wasn't convinced that regulators would dismantle Glass-Steagall well before the merger, it had to be undone to comply with Glass-Steagall.

Appleton: I agree. I think it was a foregone conclusion. I actually think that dismantling of Glass-Steagall was a good thing. The combination of market discipline on banks and the new regulatory framework will suffice to keep a lid on the overall risk of the banking industry.

Butterworth: Well, you may be right. But just because you are allowed to do something by the regulators doesn't mean you should do it. All those big plans Citigroup had after the merger of seamlessly blending banking and insurance seem to have not been realized.

Appleton: On that I have to agree with you both. The whole issue of what the boundaries of a bank should be is a fascinating one, and Citigroup's decision to sell off a part of its insurance business is an indication that banks haven't quite figured out yet what the scope of their activities should be, regardless of what regulators permit.

Moderator: I think we are all in agreement on that important point. I doubt that we'll agree on another point quite so unanimously. So, I will bring this meeting to a close.

CHAPTER ♦ 1

Basic Concepts

"Practical men, who believe themselves to be quite exempt from any intellectual influences, are usually the slaves of some defunct economist. Madmen in authority, who hear voices in the air, are distilling their frenzy from academic scribbler of a few years back. I am sure that the power of vested interests is vastly exaggerated compared with the gradual encroachment of ideas."

John Maynard Keynes: The General Theory of Employment, Interest and Money, 1947

Introduction

The modern theory of financial intermediation is based on concepts developed in financial economics. These concepts are used liberally throughout the book, so it is important to understand them well. It may not be obvious at the outset why a particular concept is needed to understand banking. For example, some may question the relevance of "market completeness" to commercial banking. Yet, this seemingly abstract concept is central to understanding financial innovation, securitization, and the off-balance sheet activities of banks. Many other concepts such as riskless arbitrage, options, market efficiency, and informational asymmetry have long shaped other subfields of finance and are transparently of great significance for a study of banking. We have thus chosen to consolidate these concepts in this chapter, to provide easy reference for those who may be unfamiliar with them.

Risk Preferences

To understand the economic behavior of individuals, it is convenient to think of an individual as being described by a utility function that summarizes preferences over different outcomes. For a wealth level W, let U(W) represent the individual's utility of that wealth. It is reasonable to suppose that this individual always prefers more wealth to less. This is called "nonsatiation" and can be expressed as $U'(W) > 0$, where the prime denotes a mathematical derivative. That is, at the margin, an additional unit of wealth always increases utility by some amount, however small.

An individual can usually be classified as being either risk neutral, risk averse or risk preferring. If *risk neutral*, the individual is indifferent between the certainty of receiving the mathematical expected value of a gamble and the uncertainty of the gamble itself. Since expected wealth is relevant for the risk neutral, and the variability of wealth is not, the utility function is *linear* in wealth, and the second derivative, denoted $U''(W)$, will equal zero. Letting $E(\bullet)$ denote the statistical expectation operator, we can write $U[E(W)] = EU(W)$ for a risk-neutral individual, where $U[E(W)]$ is the utility of the expected value of W and EU(W) is the expected utility of W. For such an individual, changing the risk of an outcome has no effect on his well-being so long as the expected outcome is left unchanged.

The utility function of a *risk-averse* individual is *concave* in wealth, that is, $U''(W) < 0$. Such an individual prefers a certain amount to a gamble with the same expected value. *Jensen's inequality* says that

$$U[E(W)] > E[U(W)]$$

if U is (strictly) concave in W. Thus, risk-averse individuals prefer less risk to more, or equivalently, they demand a premium for being exposed to risk.

A *risk–preferring* individual prefers the riskier of two outcomes having the same expected value. The utility function of a risk-preferring individual is *convex* in wealth, that is, $U''(W) > 0$, Jensen's inequality says that

$$U[E(W)] < E[U(W)]$$

if U is (strictly) convex in W.

Despite the popularity of lotteries and parimutuel betting, it is commonly assumed that individuals are risk averse. Most of finance theory is built on this assumption. Figure 1.1 depicts the different kinds of risk preferences.

In Figure 1.2 we have drawn a picture to indicate what is going on. Consider a gamble in which an individual's wealth W can be either W_1 with probability 0.5 or W_2 with probability 0.5. If the individual is risk averse, then the individual has a concave utility function that may look like the curve AB. Now, the individual's expected wealth from the gamble is $E(W) = 0.5W_1 + 0.5W_2$, which is precisely midway between W_1 and W_2. The utility derived from this expected wealth is given by U[E(W)] on the y-axis. However, if this individual accepts the gamble itself [with an expected value of E(W)], then the expected utility, EU(W), is midway between $U(W_1)$ and $U(W_2)$ on the y-axis, and can be read off the vertical axis as the point of intersection between the vertical line rising from the midpoint between W_1 and W_2 on the x-axis and the straight line connecting $U(W_1)$ and $U(W_2)$. Hence, as is clear from the picture, $U[E(W)] > EU(W)$. The more bowed or concave the individual's utility function, the more risk averse that individual will be and the larger will be the difference between U[E(W)] and EU(W).

We can also ask what sure payment we would have to offer to make this risk averse individual indifferent between that sure payment and the gamble. Such a sure payment is known as the *certainty equivalent* of the gamble. In Figure 1.2, this certainty equivalent is denoted by CE on the x-axis. Since the individual is risk

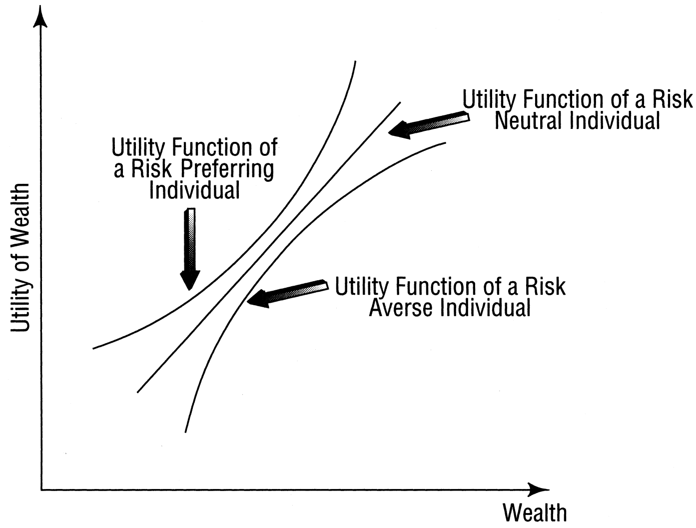


FIGURE 1.1 Three Different Types of Utility Functions

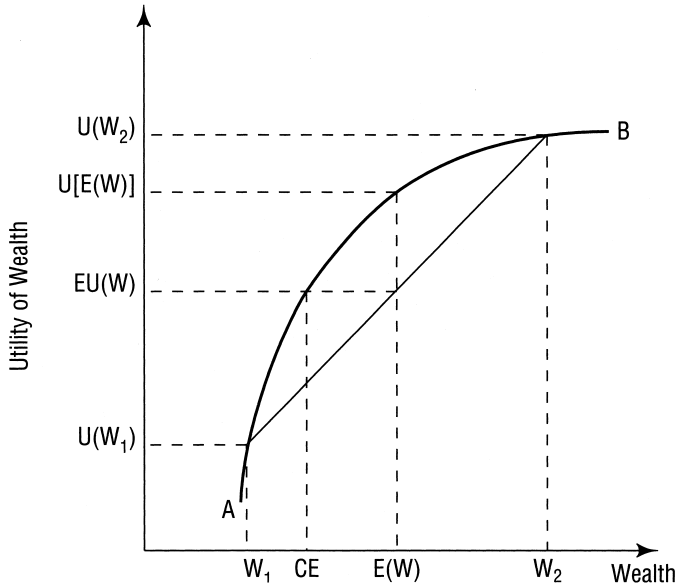


FIGURE 1.2 Risk Aversion and Certainty Equivalent

averse, the certainty equivalent of the gamble is less than the expected value. Alternatively expressed, E(W) – CE is the *risk premium* that the risk averse individual requires in order to participate in the gamble if his alternative is to receive CE for sure.

The concept of risk aversion is used frequently in this book. For example, we use it in Chapter 3 to discuss the role of financial intermediaries in the economy. Risk aversion is also important in understanding financial innovation, deposit insurance, and a host of other issues.

Diversification

We have just seen that risk-averse individuals prefer to reduce their risk. One way to reduce risk is to diversify. The basic idea behind diversification is that if you hold numerous risky assets, your return will be more predictable, but not necessarily greater. For diversification to work, it is necessary that returns on the assets in your portfolio not be perfectly and positively correlated. Indeed, if they are so correlated, the assets are identical for practical purposes so that the opportunity to diversify is defeated. Note that risk can be classified as idiosyncratic or systematic. An idiosyncratic risk is one that stems from forces specific to the asset in question, whereas systematic risk arises from the correlation of the asset's payoff to economy-wide phenomena such as depression. Idiosyncratic risks are diversifiable, systematic risks are not.

To see how diversification works, suppose that you hold two assets, A and B, whose returns are random variables.[1] Let the variances of these returns be σ_A^2 and σ_B^2, respectively. Suppose the returns on A and B are perfectly and positively correlated, so that $\rho_{AB} = 1$, where ρ_{AB} is the correlation coefficient between A and B. The proportions of the portfolio's value invested in A and B are y_A and y_B, respectively. Then the variance of the portfolio return is

$$\sigma_p^2 = y_A^2\sigma_A^2 + 2y_Ay_B\mathrm{Cov(A,B)} + y_B^2\sigma_B^2 \qquad [1.1]$$

where Cov(A,B) is the covariance between the returns on A and B. Then, using

$$\mathrm{Cov(A,B)} = \rho_{AB}\sigma_A\sigma_B \qquad [1.2]$$

we have

$$\sigma_P^2 = y_A^2\sigma_A^2 + 2y_Ay_B\rho_{AB}\sigma_A\sigma_B + y_B^2\sigma_B^2 \qquad [1.3]$$

Since $\rho_{AB} = 1$, the right-hand size of (1.3) is a perfect square, $(y_A\sigma_A + y_B\sigma_B)^2$. As long as $y_A\sigma_A + y_B\sigma_B \geq 0$, we can write (1.3) as

$$\sigma_p = y_A\sigma_A + y_B\sigma_B. \qquad [1.4]$$

Thus, if $\rho_{AB} = 1$, the standard deviation of the portfolio return is just the weighted average of the standard deviations of the returns on assets A and B. *Diversification therefore does not reduce portfolio risk when returns are perfectly and positively*

1. Suppose x and z are two random variables that can each take any value from $-\infty$ to $+\infty$. A random variable is one whose behavior is described by a probability density function, but its precise value is unknown. Let f(x) and g(z) be the density functions of x and z, respectively. Then, the probability that x will lie between the two numbers a and b is $\Pr(a \leq x \leq b) = \int_a^b f(x)dx \geq 0$, and $\int_{-\infty}^{\infty} f(x)dx = 1$. The statistical mean (expected value) of x is $E(x) = \int_{-\infty}^{\infty} xf(x)dx$, its variance is $\sigma_x^2 = \int_{-\infty}^{\infty} [x - E(x)]^2 f(x)dx$, and the mean and variance of z are analogously defined. The covariance of x and z is $\mathrm{Cov}(x,z) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [x - E(x)][z - E(z)]\, f(x)g(z)dxdz$ and the correlation between x and z is $\rho_{xz} = \mathrm{Cov}(x,z)/\sigma_x\sigma_z$ where σ_x and σ_z are the standards deviations (square roots of the respective variances) of x and z, respectively.

correlated. For any general correlation coefficient ρ_{AB}, we can write the portfolio return variance as

$$\sigma_p^2 = y_A^2\sigma_A^2 + 2y_A y_B \rho_{AB}\sigma_A\sigma_B + y_B^2\sigma_B^2. \quad [1.5]$$

Holding fixed y_A, y_B, σ_A and σ_B, we see that $\partial\sigma_p^2/\partial\rho_{AB} > 0$, that is, portfolio risk increases with the correlation between the returns on the component assets. At $\rho_{AB} = 0$ (uncorrelated returns),

$$\sigma_p^2 = y_A^2\sigma_A^2 + y_B^2\sigma_B^2. \quad [1.6]$$

Example 1.1 To see that diversification helps in this case, suppose $y_A = y_B = 0.5$, $\sigma_A^2 = 100$, $\sigma_B^2 = 144$. Calculate the variance of a portfolio of assets A and B, assuming first that the returns of the individual assets are perfectly positively correlated, $\rho_{AB} = 1$, and then that they are uncorrelated, $\rho_{AB} = 0$.

Solution In the case of perfectly and positively correlated returns, $\sigma_P = 0.5(10) + 0.5(12) = 11$, or $\sigma_p^2 = 121$. With uncorrelated return, (1.6) implies that $\sigma_p^2 = 0.25(100) + 0.25(144) = 61$. Thus, not only is this variance lower than with perfectly and positively correlated returns, but it is also lower than the variance on either of the components assets.

The maximum effect of diversification occurs when ρ_{AB} is at its minimum value of -1, that is, returns are perfectly negatively correlated. In this case

$$\sigma_p^2 = y_A^2\sigma_A^2 - 2y_A y_B\sigma_A\sigma_B + y_B^2\sigma_B^2 \quad [1.7]$$

so that

$$\sigma_p = |y_B\sigma_B - y_A\sigma_A|. \quad [1.8]$$

This seems to indicate that the portfolio will have some risk, albeit lower than in the previous cases. But suppose we construct the portfolio so that the proportionate holdings of the assets are inversely related to their relative risks. That is,

$$y_A/y_B = \sigma_B/\sigma_A \quad [1.9]$$

or

$$y_A = \sigma_B y_B/\sigma_A. \quad [1.10]$$

Substituting (1.10) in (1.8) yields

$$\sigma_p = y_B\sigma_B - (\sigma_B y_B\sigma_A/\sigma_A) = 0$$

indicating that in this special case of perfectly negatively correlated returns, portfolio risk can be reduced to zero!

Even when assets with perfectly negatively correlated returns are unavailable, we can reduce portfolio risk by adding more assets (provided they are not perfectly positively correlated with those already in the portfolio), while keeping fixed the total wealth invested in the portfolio.[2] To illustrate, suppose we have N assets available, each with returns pairwise uncorrelated with the returns of every other asset. In this case, a generalized version of (1.6) is

$$\sigma_p^2 = \sum_{i=1}^{N} y_i^2 \sigma_i^2 \qquad [1.11]$$

where y_i is the fraction of the portfolio value invested in asset i, where $i = 1, \ldots, N$, and σ_i^2 is the variance of asset i. Suppose we choose $y_i = 1/N$.

Then, defining σ_{max}^2 as the maximum variance among the σ_i^2 (we assume $\sigma_{max}^2 < \infty$, and permit $\sigma_i^2 = \sigma^2$ for all i in which case $\sigma_{max}^2 = \sigma^2$), (1.11) becomes

$$\begin{aligned} \sigma_p^2 &= \sum_{i=1}^{N} \left[\frac{1}{N}\right]^2 \sigma_i^2 \\ &\leq N\left[\frac{1}{N}\right]^2 \sigma_{max}^2 \\ &= \frac{\sigma_{max}^2}{N}. \end{aligned}$$

As N increases, σ_p^2 diminishes, and, in the limit, as N goes to infinity, σ_p^2 goes to zero. Thus, if we have sufficiently many assets with (pairwise) uncorrelated returns, we can drive portfolio risk as low as we wish and make returns as predictable as desired.

An obvious question is why investors do not drive their risks to zero. First, not all risks are diversifiable. Some contingencies affect all assets alike and consequently holding more assets will not alter the underlying uncertainty. This is the notion of *force majeure* in insurance. Natural calamities such as floods and earthquakes are examples, as are losses attributed to wars. Second, as the investor increases the number of securities held in the portfolio, there are obvious costs of administration. These costs restrain diversification, but in addition numerous studies indicate that a large fraction of the potential benefits of diversification are obtained by holding a relatively small number of securities. That is, the marginal benefits of diversification decline rapidly as the number of securities increases.

Finally, cross-sectional reusability of information diminishes the incentive to diversify. We shall have more to say in Chapter 3 about information reusability since this is a major motivation for the emergence of financial intermediaries. Suffice to say that if a lender invests to learn about a customer in the steel business in order to make a loan, it will see a potential benefit to lending to others in the steel business. The resulting concentration spreads the costs of becoming informed. Thus, we observe diversification within areas of specialization among most financial intermediaries.

2. In pointing out the "fallacy of large numbers," Samuelson (1963) shows that diversification is not necessarily preferred by a risk-averse individual if one also adds more wealth to the portfolio as more assets are added. We will have more to say about this in Chapter 3.

And when we speak of financial intermediaries processing risk, we mean that they are typically diversifying some, absorbing some, and shifting some to others.

The concept of diversification is used in this book in a variety of different contexts. We use it quite extensively in Chapter 3, for example, to explain economies of scale in the production of financial intermediation services.

Riskless Arbitrage

Arbitrage is the simultaneous purchase and sale of identical goods or securities that are trading at disparate prices. This opportunity for riskless profit is transitory because the exploitation of such opportunities eliminates the initial price disparities.

The term *arbitrage* is often loosely applied to situations in which objects of trade are similar, but not identical, and where the risk is thought to be small but not totally absent. Since such situations are often referred to as arbitrage, the redundant ‘‘riskless arbitrage’’ has emerged to describe arbitrage rather than limited risk speculation (a situation in which a profit can be had for a small risk). Thus, succinctly defined, riskless arbitrage is profit without risk and without investment. We shall later discuss ‘‘risk-controlled arbitrage’’ as an illustration of limited risk speculation. Consider the following illustration of riskless arbitrage.

Example 1.2 Suppose that there are two possible states of the economy next period: high (H) and low (L). Available in the capital market are two risky securities, R_1 and R_2, and a riskless bond, B. The state-contingent payoffs and current market prices of these instruments are presented in Table 1.1 below. Examine whether there are riskless arbitrage opportunites.

TABLE 1.1 State-Contingent Payoffs and Prices of Securities

	Payoff in State		
Security	**H**	**L**	**Current Price**
R_1	\$100	0	\$40
R_2	0	\$100	\$40
B	\$50	\$50	\$43

Solution Since you can combine R_1 and R_2 to get a payoff that is equivalent to that from B, you can see now that there is an opportunity for riskless arbitrage. If you buy one unit each of R_1 and R_2 for a total outlay of \$80, you are assured of \$100 next period, regardless of whether state H or L is realized. So you can sell two units of B for \$86 earning a riskless profit of \$6. You are obliged to pay the buyers of these two units of B a total of \$100 next period, but this you can do from the cash inflows produced by the R_1 and R_2 that you possess. Since you can sell these two units of B before you even buy R_1 and R_2, your profit requires no investment on your part and no risk. You could of course sell an arbitrarily large number of units of B and buy the appropriate units of R_1 and R_2, giving yourself a veritable money machine. But as your purchases and sales

(*Continued*)

increase in volume, it is reasonable to expect the prices of the securities to converge, thereby eliminating the opportunity for riskless arbitrage again. An important implication is that the prices of related securities cannot be determined independently of each other. This observation has provided a powerful way to price derivative securities such as options.

The notion that any capital market equilibrium should preclude riskless arbitrage has proved to be a powerful concept in many applications in finance, including financial intermediation. We will see this idea applied in other contexts, including the valuation of contingent claims such as loan commitments.

Options

An option is a contract that gives the owner the right to either buy or sell an asset at a predetermined price at some future time or over some fixed time interval. Consider an asset whose value at time $t = 1$ will be X. Viewed at $t = 0$ (the present), X is a random variable. A *call option* entitles its owner to *buy* this asset at a fixed price, P_c, at or before $t = 1$. If he does not wish to buy the asset, he can allow the option to expire unexercised. Thus, the value of the call option at $t = 1$ is

$$C(t = 1) = \begin{cases} X - P_c & \text{if } X > P_c \\ 0 & \text{if } X \leq P_c. \end{cases} \quad [1.12]$$

The theory of option pricing explains $C(t = 0)$, the value of the call option at $t = 0$. The basic idea is to construct a portfolio consisting of the underlying stock and a riskless bond in such a manner that it yields the same payoff as the option. Since there can be no riskless arbitrage in equilibrium, the prices of this portfolio should equal the price of the option. We can then price the option by using the observed prices of the stock and the bond. We will have more to say about option pricing in later chapters.

Symmetrically, a *put option* entitles the option owner to *sell* an asset at a fixed price, P_p, at or before $t = 1$. Thus, at $t = 1$ the value of the put option is

$$P(t = 1) = \begin{cases} P_p - X & \text{if } X < P_p \\ 0 & \text{if } X \geq P_p. \end{cases} \quad [1.13]$$

In addition to being a put or call, an option can be either *European* or *American.* A European option can be exercised only at some predetermined maturity date, for example, at $t = 1$ in the above discussion. An American option can be exercised any time prior to maturity. Thus, an American option never can be worth less than its European counterpart.

An important property of options that we will use frequently is *that the more volatile the value of the underlying security on which the option is written, the more valuable the option*. The following example illustrates this property.

Example 1.3 Consider a European call option with an exercise price $P_c = \$100$. At $t = 1$, X will be \$110 with probability 0.5 and \$90 with probability 0.5. For simplicity, suppose everybody is risk neutral and the discount rate is zero (so that future payoffs are valued the same as current payoffs). Then from (1.12) we have

$$C(t = 1) = \begin{cases} \$10 & \text{with probability } 0.5 \\ 0 & \text{with probability } 0.5 \end{cases}$$

Thus, $C(t = 0) = 0.5(10) = \$5$. Now suppose we increase the variance of X, keeping its mean unchanged. Let X be \$150 with probability 0.5 and \$50 with probability 0.5. From (1.13) we have

$$C(t = 1) = \begin{cases} \$50 & \text{with probability } 0.5 \\ 0 & \text{with probability } 0.5 \end{cases}$$

Thus, $C(t = 0) = 0.5(50) = \$25$. The call option is now five times more valuable! You should work through a similar example for put options to convince yourself that puts have the same property.

Option pricing theory is used in our later discussions of the valuation of off-balance sheet claims like loan commitments, and in our analysis of deposit insurance.

Market Efficiency

An efficient capital market is one in which every security's price equals its "true" economic value. But what is true? In economics, it means a price that incorporates *all* the information available to investors at the time. In an efficient market, an *appropriately defined set of information* is fully and immediately impounded in the prices of all securities. The basic idea is that competition among investors and the resulting informational exchanges will lead to market efficiency. This implies that price changes in an efficient market must be random. If prices always reflect all relevant information, then they will change only when new information arrives. However, by definition, new information cannot be known in advance. Therefore, price changes cannot be predictable.

We speak of three forms of market efficiency, distinguished by the amount of information impounded in the price. A market is said to be *weak-form* efficient if prices impound all historical information. In a weak-form efficient market, if P_t is the price at time t, then the expected value (at time t) of the price at time $t + 1$ conditional on the price at time t, written as $E(P_{t+1}|P_t)$, is the same as $E(P_{t+1}|P_t, \ldots, P_0)$, the expected value of P_{t+1} conditional on the entire history of stock prices up until time t (that is, $P_t, \ldots, P_0$). That is,

$$E(P_{t+1}|P_t) = E(P_{t+1}|P_t, P_{t-1}, P_{t-2}, \ldots, P_0). \quad [1.14]$$

This means that you can do no better forecasting tomorrow's price P_{t+1} using the entire history of prices than you could using just today's price P_t. The reason is that weak-form efficiency implies that P_t itself should contain all the historical information contained in the sequence $\{P_{t-1}|P_{t-2}, \ldots, P_0\}$.

Semistrong form market efficiency requires that all publicly available information be contained in the current price. Since all historical information is in the public domain, a semistrong form efficient market is always weak-form efficient. However, there may be contemporaneous information in the public domain that became available after the most recent price was determined. Thus, semistrong form efficiency is a more demanding form of efficiency than weak-form efficiency.

A market is *strong-form* efficient if prices impound *all* information, including that possessed by insiders. Few economists believe that markets are strong-form efficient. Although there is a mountain of empirical evidence accumulated over nearly 2 decades suggesting that markets are semistrong form efficient, recent theoretical and empirical research has shown that the market may not even be weak-form efficient.[3]

If the capital market were strong-form efficient, there would be no role for financial intermediaries as information processors (unless intermediaries were crucial in making the market efficient). However, when strong-form efficiency fails to obtain, we can have different individuals primarily possessing different sorts of information. In Chapter 3 we will show that in such markets, financial intermediaries have a role to play. At many junctures in this book, we will discuss how the efficiency (or lack thereof) of markets affects the profits to be earned from financial intermediation. An example of this is financial innovation.

Market Completeness

The economic world we inhabit is complex and pervasively uncertain. It is often useful to think of this uncertainty in terms of the possible states of nature that can occur in the future. Each such state, call it θ, can be viewed as a possible economic outcome. For example, θ may correspond to different levels of gross domestic product. Although we do not know what θ will be tomorrow, we can assign a probability distribution over possible values of θ. For the theory, it does not matter how many values θ can take. For simplicity, suppose θ can take integer values from 1 to some arbitrary number N.

In evaluating problems of economic efficiency, an important consideration is the *number* of different *financial securities* available relative to the number of states of nature. Two financial securities are considered "different" if they do not have identical payoffs in every state. To see the implications of this, consider the following simple example.

Example 1.4 Suppose there are three states of nature and only two securities may be thought of as shares of stock issued by two different companies. The payoffs offered by these securities in the different states of nature are shown in Table 1.2.

3. See Brown and Jennings (1990) and Lehmann (1990), for example.

TABLE 1.2 Example With Three States of Nature and Two Securities

	States of Nature		
	1	**2**	**3**
Security 1 payoffs	10	20	15
Security 2 payoffs	15	0	25

Consider now an individual who owns 10 percent of security 1 and 20 percent of security 2. If $\theta = 1$ occurs, his wealth will be $0.10(10) + 0.20(15) = 4$. If $\theta = 2$ occurs, his wealth will be $0.10(20) + 0.20(0) = 2$. If $\theta = 3$ occurs, his wealth will be $0.10(15) + 0.20(25) = 6.5$. Thus, the value of the individual's portfolio can be described by the vector (4, 2, 6.5), where the first element corresponds to his wealth in state 1 and so on.

While the individual can achieve the vector (4, 2, 6.5), it is easy to see that one cannot achieve the vector (2, 6.5, 9.5). It is impossible for one to find ownership fractions in the two securities that will allow one to achieve this wealth vector. The reason is that there are fewer (independent) securities than there are states of nature. If we had a third security, we could have ensured that our individual could achieve any desired income vector. Of course, in reality individuals are also constrained by their budgets. The point is simply that when there are fewer securities than there are states of nature, it is generally impossible for the individual to attain any desired future wealth rearrangement. This is ultimately a limitation on the individual's ability to insure against contingencies.

The securities depicted in our simple example are not really stocks or bonds or any of the other financial securities commonly found in the capital market. Rather, these securities are claims to income in different states of the world. We can nevertheless visualize a market where such claims are traded. We would then have a number of securities, one for each state of nature, promising to pay 1 dollar if that particular state occurred and nothing otherwise. Such securities are called *primitive state-contingent claims* or *Arrow-Debreu securities* after the economists Kenneth Arrow and Gerard Debreu, who first studied this issue and later went on to win Nobel Prizes in Economics. Such a market would represent an ideal way of organizing a securities exchange, since it would give individuals complete freedom (subject only to their own purchasing power limitations) in designing portfolios that deliver the desired distribution of income in different states of the world. That is, an individual can design any "homemade" security in such a market.

If there are as many Arrow-Debreu securities as there are states of nature, the market is referred to as *complete*. In a complete market, an individual can achieve any desired distribution of income, subject to the individual's budget constraint. On the other hand, if there are fewer Arrow-Debreu securities than there are states of nature, we have an *incomplete market*, which places a limitation on the ability of transactors to manage uncertainty. The conceptual beauty of a complete market is that we can examine the market prices of securities that are currently trading and determine the market price of any *new* security we may wish to introduce. We can do this without knowing the preferences of individual investors in the economy. The key is that we can use the prices of existing securities to compute the prices of the (fictitious) Arrow-Debreu securities, and then use this information to price any new security we want to introduce. Suppose that in Example 1.2, we are given the prices of securities R_1 and R_2; recall that the price of each security is \$40. Moreover, R_1 pays off \$100 in state H and 0 in state L, whereas R_2 pays off 0 in state H and \$100 in state L. Let Pi_H and Pi_L be the prices of the Arrow-Debreu securities in states H and L, respectively. Then, the market price of security R_1 should be 100 times the price of the state H Arrow-Debreu claim, that is, $40 = 100\ Pi_H$, $Pi_H = 0.4$. Similarly, the market price of security R_2 should be 100 times the price of the state L Arrow-Debreu claim, that is, $Pi_L = 0.4$. We are now ready to price any security in this two-state economy. For example, the riskless bond in example 1.2, which pays \$50 in each state, should be priced at $50Pi_H + 50Pi_L = \$40$. A security that pays \$1,000 in state H and \$56 in state L should sell at $1000Pi_H + 56Pi_L = \$422.40$, and so on.

The concept of market incompleteness is used in Chapter 12 in connection with our discussion of financial innovation. Other applications can be found in chapters on off-balance sheet activities, securitization, and deposit insurance.

Asymmetric Information and Signaling

Economic transactions often involve people with different information. For example, the borrower usually knows more about its own investment opportunities than the lender does. Corporate insiders normally know more about the values of assets owned by their firms than shareholders. A doctor can be expected to be better informed about his or her own medical expertise than a patient.

The better informed economic agents have a natural incentive to exploit their informational advantage. Insider trading scandals on Wall Street illustrate how those with access to privileged information can profit, despite laws aimed at preventing such activity. Of course, those who are uninformed should anticipate their informational handicap and behave accordingly. It is this interaction between the inclination of the informed to strategically manipulate and the anticipation of such manipulation by the uninformed that results in distortions away from the "first best" (the economic outcome in a setting in which all are equally well-informed).

Problems of asymmetric information were brought to the forefront when George Akerlof (1970), who later went on to win the Nobel Prize in Economics for his contribution, sought to explain why used cars sell at such large discounts relative to the prices of new cars. The following example takes some shortcuts, but conveys the intuition of Akerlof's analysis.

Example 1.5 Consider a used car market in which differences in the care with which owners use their cars lead to quality differences among cars that started out identical. It is natural to suppose that the owner of the used car knows more about its quality than potential buyers. As an example, assume that there are three possible quality levels that the used car in question can have, $q_1 > q_2 > q_3 = 0$. If the quality level is q_3, the car is a lemon. Such a car would be priced as being worthless if buyers could correctly assess its quality. If the quality is q_2, the car has a value of \$5, and if the quality is q_1, the car is worth \$10. Assume that all agents are risk neutral and a buyer does not want to pay more for a car than its expected worth. In like vein, the car owner does not wish to sell at less than what the car is worth. Suppose that each car owner knows his car's quality, but buyers only know that cars for sale can be of quality q_1, q_2, or q_3. Faced with a given car, they cannot identify its precise quality. However, they believe that there is a probability 0.4 that the quality is q_1, a probability 0.2 that it is q_2, and a probability 0.4 that it is q_3. What will happen in such a market?

Solution If all cars are offered for sale, risk neutral buyers will compute the *expected* value of a (randomly chosen) car as $(0.4) \times \$10 + (0.2) \times \$5 + (0.4) \times 0 = \$5$. Hence, if the market is competitive, we would expect \$5 to be the market clearing price. However, at this price those who own cars with quality q_1 will refuse to sell. Thus, only cars of qualities q_2 and q_3 will be offered at \$5. However, buyers will anticipate this and revise their beliefs about the quality dispersion of cars in the market. They will now assume that if the selling price is \$5, the probability is $0.2/(0.2 + 0.4) = 1/3$ that the quality is q, and it is 2/3 that it is q_3. Thus, the expected value of a car drops to $(1/3)(5) + (2/3)(0) = \$1.67$. No cars will, therefore, be bought at \$5 (it cannot be a market clearing price). Now if \$1.67 is the price, those with cars of quality q_2 will drop out and the only cars offered for sale will be lemons. This process is called *adverse selection* and it results in the market clearing price being driven to zero. In other words, the demand for cars at any positive price is zero, and the market breaks down, as depicted in Figure 1.3. You should note a key assumption made in this example. All market participants have *rational expectations*. That is, uninformed buyers rationally anticipate what informed sellers will do at any given price and informed sellers rationally anticipate the demand buyers will have at that price. Hence, we don't need to go through a sequential process of price convergence to zero. No cars will be bought or sold.

The insight that asymmetric information can cause market failure was novel and striking. Its profound implications were quickly recognized to extend well beyond the used car market. Informational asymmetries were seen as being capable of causing markets to break down and thus possibly justify regulatory intervention by the government. Indeed, in the chapters that follow, we will examine banking regulation from this informational perspective.

However, calls for regulation based on Akerlof's analysis were too hasty. Market participants have the capability and incentives to deploy mechanisms to prevent market failure, and in any case market failure is the most extreme form of distortion created by asymmetric information. To see this in the context of our used car example, consider the following extension of that example.

Example 1.6 Suppose that cars of different qualities have different probabilities of engine failure within a given time period, and that these differences are reflected in their values of 0, \$5 and \$10. Suppose the failure probability is 0.1 for the q_1 quality car, 0.5 for the q_2 quality car, and 1 for the q_3 quality car. Do warranties have a role to play in this market?

Solution To prevent market failure, the sellers of better cars must somehow distinguish themselves from the sellers of lower quality cars. One way to do this would be

(*Continued on page* 27)

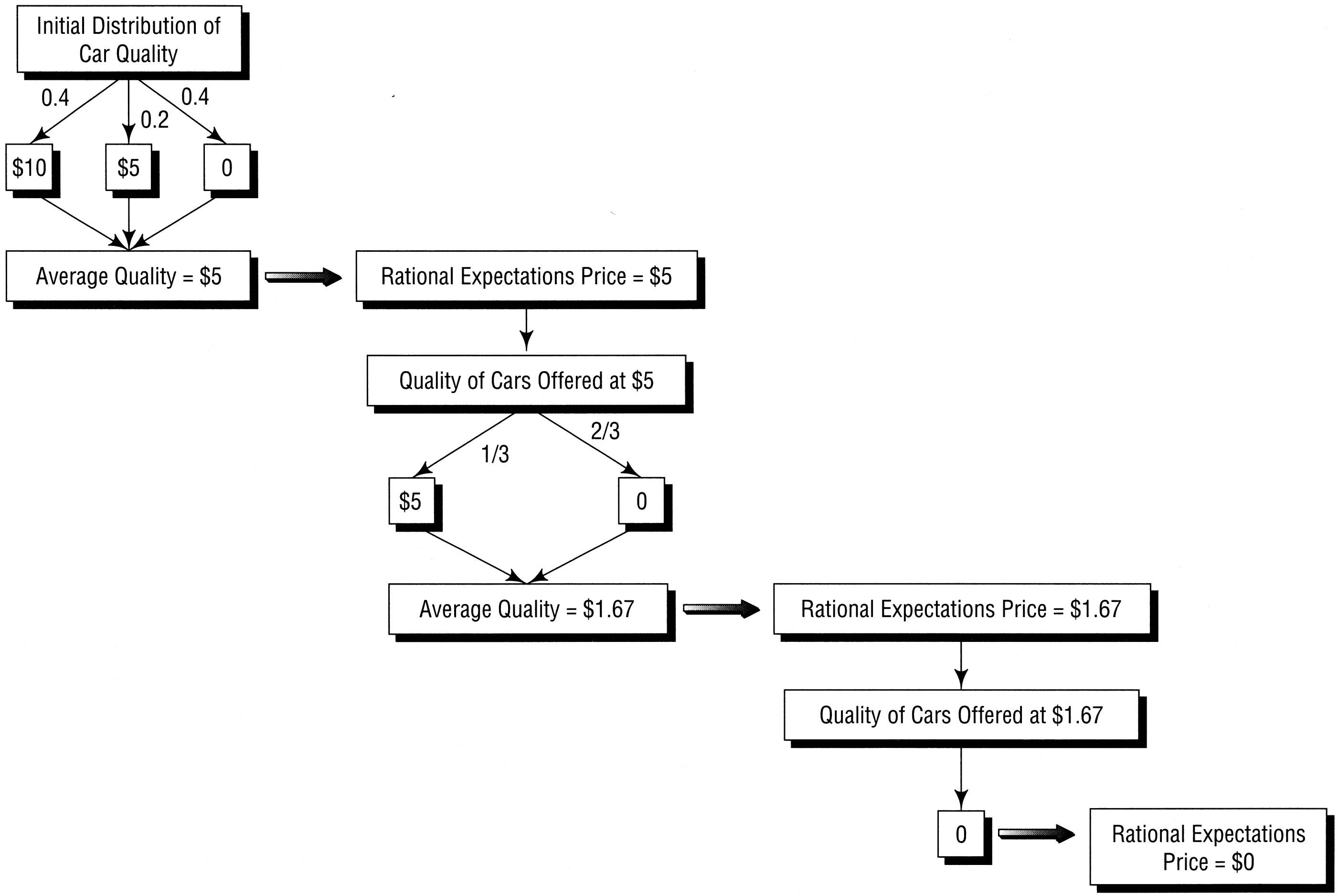


FIGURE 1.3 A Pictorial Depiction of the Adverse Selection Process

with warranties or guarantees. The seller of the q_1 quality car can announce that he will reimburse the buyer $\$W_1$ if his car fails, and the seller of q_2 quality car can announce that he will pay the buyer $\$W_1$ if his car fails. If buyers believe that *only* the owners of q_1 quality cars will promise a $\$W_2$ payment upon failure and that only the owners of q_2 quality cars will promise a $\$W_2$ payment upon failure, then they will make the appropriate inference and should be willing to pay prices that accurately reflect the qualities of the cars offered for sale. In order for such an indirect transfer of information to be effective, no seller should wish to mimic the strategy of a seller of a different quality car. Otherwise, buyers will eventually learn of the potential mimicry and the credibility of the signal will be destroyed.

Since the failure probability for a q_1 quality car is 0.1, the buyer should be willing to pay \$10 (the intrinsic worth of a q_1 quality car) plus 0.1 times W_1, the latter being the amount he expects to collect from the seller. Thus, the equilibrium price (P_1) of a q_1 quality car should be $\$10 + 0.1W_1$. Similarly, if the owner of a q_2 quality car follows his equilibrium strategy, the equilibrium price (P_2) of a q_2 quality car should be $\$5 + 0.5W_2$. To ensure that the q_2 quality car owner will not misrepresent himself as a q_1 quality car owner, W_1 should be set to satisfy

$$10 + 0.1W_1 - 0.5W_1 \leq 5 + 0.5W_2 - 0.5W_2. \qquad [1.15]$$

The left-hand side (LHS) of (1.15) is the expected payoff to a q_2 quality car owner misrepresenting himself as a q_1 quality car owner; he receives a price P_1 and has an expected outflow of 0.5 W_1 to pay the liability under the warranty. The right-hand side (RHS) of (1.15) is what the q_2 quality car owner gets if he follows his nonmimic strategy; he receives a price of P_2 and has an expected cash outflow of $0.5W_2$. When someone is indifferent between telling the truth and lying, it is conventionally assumed that truth-telling will be chosen. Thus, (1.15), which is referred to as an *incentive compatibility* (*IC*) *condition*, can be treated as an equality and we can solve it to obtain $W_1 = 12.5$. Incentive compatibility here means that the seller's incentives to maximize personal profit should be compatible with truthful representation of the car's quality.

The IC condition that ensures that the seller of lemons does not mimic the seller of q_2 quality cars can be similarly expressed as follows

$$5 + 0.5W_2 - W_2 \leq 0. \qquad [1.16]$$

Solving (1.16) as an equality yields $W_2 = 10$. It is straightforward to verify that the seller of q_2 quality cars will not mimic the seller of lemons under the described conditions, that is, q_2 quality cars will be offered for sale.

You can easily verify that this scheme guarantees that the seller of lemons will not mimic the seller of q_1 quality cars and that the seller of q_1 quality cars will not mimic either the seller of q_2 quality cars or the seller of lemons.

To summarize, we have produced a simple scheme of "warranties" that prevents market failure. The seller of q_1 quality cars promises to pay the buyer \$12.5 if his car fails; this enables him to sell his car for $10 + 0.1(12.5) = \$11.25$. The seller of q_2 quality cars promises to pay the buyer \$10 if his car fails; this enables him to sell his car for $5 + 0.5(10) = \$10$. The lemons are withdrawn from the market.

The warranty offered here can be viewed as a *signal* of quality. A (perfectly revealing) signal is one that enables the uninformed to infer which the informed agent knew privately *a priori*. For a signal to be useful it must be *informative*, and this requires that the signaling mechanism be incentive compatible. In turn, incentive compatibility requires that the cost of signaling must be negatively correlated with quality,[4] Michael Spence too was awarded the Nobel Prize in Economics for his contribution to the economics of asymmetric information. That is, it must be less costly at the margin for a higher quality seller to emit a given signal. The higher cost of signaling serves to deter the lower quality sellers from mimicking their higher quality counterparts. In our context, you can see that a warranty of \$12.50 imposes an expected liability of \$1.25 on the q_1 quality seller, \$6.25 on the q_2 quality seller and \$12.50 on the seller of lemons.

Note too that *in equilibrium* (that is, when each seller maximizes expected profit) the chosen signal is *costless* for the seller emitting it. Although the q_1 quality seller promises to pay \$12.50, he has only a 0.1 probability of having to pay, and since he collects \$11.25 upon selling the car, his cash inflow net of the expected liability is \$10. This is exactly what he'd have gotten *without* issuing a warranty, if we were in a "first best" world in which the quality of each car was common knowledge. Likewise, the q_2 quality seller's net cash inflow is \$5. Signals are *costless* in equilibrium. The reason for this, as you may have guessed, is that the seller is (correctly) compensated by the buyer for issuing the warranty, that is, cars with better warranties sell at higher prices. Such signals are called *nondissipative*[5] because the cost of the signal is a *transfer payment* from one party to the other, and there is no loss in the aggregate.

We can also have *dissipative* signals. To see this, suppose that instead of paying cash, the seller promises to reimburse the cost of repairing a portion of the damage. The q_1 quality seller promises complete coverage, the q_2 quality seller offers to absorb half the cost of repair, and the lemons owners choose not to participate. For every dollar it costs the seller to fix the damage, its value in terms of improved car quality is \$0.80. You can now easily verify that there exists a signaling scheme similar to the one derived previously that ensures truthful signaling by each seller, assuming that the seller is willing to accept a net payoff (after dissipative signaling costs are deducted) that is less than the car's worth. The q_1 quality seller's net receipt is less than \$10 and the q_2 quality seller's less than \$5. Each absorbs a signaling cost for which it is not compensated, that is, there is a net loss due to signaling.[6] For example, dividends can be a dissipative signal of future cash flows if they are personally taxed at a higher rate than capital gains (as was the case prior to the 1986 Tax Reform Act) and if external financing involves (transactions) costs that are avoided by financing with retained earnings.[7] Later in this book we will see other examples of dissipative signaling.

The concept of asymmetric information underlies much of what we discuss in this book, so you should expect to encounter it in more than a few of the remaining chapters.

4. See Spence (1973, 1974). Michael Spence too was awarded the Nobel Prize in Economics for his contributions to the economics of asymmetric information.

5. See Bhattacharya (1980).

6. If the seller is unwilling to bear the dissipative cost of signaling and the buyer will not bear it either, then a signaling equilibrium will fail to exist.

7. See, for example, Bhattacharya (1979).

Agency and Moral Hazard

It has been observed that the key distinction between man and machine is *moral hazard*.[8] First introduced in the insurance literature, this term describes situations in which the incentives of principal (the employer or the owner of the property) and agent (the employee or the person renting/using the property) diverge. A rational economic agent can be expected to maximize his own expected utility,[9] and where his self-interest conflicts with the principal's, the principal will suffer. The principal must therefore design a contract that will achieve a congruence between her goals and the agent's.

Examples of moral hazard abound. Consider automobile insurance. If you have a car that you know is worth $500 and your collision insurance will pay you $1,000 if the car is completely destroyed, you may be tempted to let your car roll down the hill and collide with an immovable object. Now you may never dream of doing this, but your willingness to spend on the maintenance of brakes may be subtly affected by your insurance policy. In any case, insurance companies cannot afford to assume that ethical or reputational considerations dominate their customers' behaviors.[10] This is one reason why we observe deductibles in insurance contracts. Coinsurance clauses are designed to share the risks and thereby bring the insured's incentives into closer alignment with those of the insurer.[11]

Moral hazard is also common in financial contracting among claimants in a corporation. Suppose you manage a firm and your goal is to maximize shareholder wealth. If you have risky bonds outstanding, you will not always choose investments that maximize the total value of the firm. Rather, you may choose projects that maximize the value of equity at the expense of the bondholders. This can be illustrated with the following numerical example.

8. Ross (1974).
9. We will refer to agents in the masculine and principals in the feminine.
10. Reputation enters via the customer's concern regarding future insurance premiums.
11. An interesting illustration of moral hazard is provided by the following report in *The Wall Street Journal* (WSJ) of October 10, 1990, titled, "More Car Owners are Scheming to Cheat Insurance Companies as Economy Falters."

"When a popular Dallas-area swimming hole developed a mysterious oil slick two months ago, it didn't take police long to discover something fishy was going on.

Littering the bottom of the abandoned stone quarry were 20 late-model automobiles, including a mint-condition 1990 Chevrolet Blazer. All of them had been reported stolen, and insurance companies had already paid off the owners. But contrary to claims in reports filed with insurance companies, most of the cars had keys in the ignition. And none of the vehicles had been stripped of fancy stereos, wheels or other easy to get accessories.

The police conclusion: The cars weren't stolen at all but had been dumped by their owners in what investigators say is one of the biggest 'car dunking' insurance scams in Texas history.

Hard figures aren't available, but most experts say 10% to 15% of all claim dollars paid out on car insurance result from some form of fakery. According to the Insurance Information Institute, that works out to between $5.4 billion and $8.1 billion of the $54 billion in claims paid last year."

—Michael Allen, Staff Reporter of the *WSJ*

Example 1.7 Consider a firm that will liquidate one period hence at time t = 1. There are no taxes and the firm can invest \$30 in a risky venture at t = 0 using retained earnings. If the investment is not made, shareholders get a dividend of \$100 at t = 0. The firm's debt requires a payment of \$100 at t = 1, and its investment choices are described in Table 1.3.

TABLE 1.3 Payoffs Related to Different Investment Opportunities

	State of Nature	
Strategy	**Boom (with probability 0.5)**	**Bust (with probability 0.5)**
Total firm value at $t = 1$ if no investment made and \$100 dividend paid at $t = 0$	\$110	\$70
Total firm value at $t = 1$ if \$30 investment made and \$70 dividend paid at $t = 0$	\$200	\$ 5

For simplicity, assume that the discount rate is zero. What should the firm do?

Solution To analyze this problem, first compute the net present value (NPV) of each choice for the firm as a whole. If it does not invest, then its expected value is 0.5(110) + 0.5(70) = \$90. Add to this the \$100 dividend paid at t = 0 and we get a total firm value of \$190. If it does invest, then its expected value is 0.5(200) + 0.5(5) = \$102.5. Add to this the \$70 dividend paid at t = 0 and we get a total firm value of \$172.5. Since total firm value is lower with the investment than without, the project has negative NPV. The apparent choice should be to reject the investment.

Hold it for a minute, though! This decision rule is the right one only if you want to maximize total firm value. But remember that your goal is to maximize the wealth of the shareholders. If there is no investment, the shareholders get \$100 dividend plus \$10 (\$110 debt payment) in the boom state and nothing in bust state (limited liability, which stipulates that the liability of the shareholder does not extend beyond the assets of the firm, means that the bondholders get \$70 and the shareholders get 100 + 0.5(10) = \$105. On the other hand, if the project is accepted, they get \$100 in the boom state and nothing in the bust state. Thus the value of this strategy to the shareholder is 70 + 0.5(100) = \$120. Clearly, the shareholders want you to invest in the project. Thus, a project with negative NPV for the firm as a whole may be chosen in the best interest of the shareholder.

This example illustrates a moral hazard faced by bondholders. The firm, acting in the interest of the shareholders, has an incentive to undertake investments that benefit the shareholders at the expense of creditors. In this example, the expected payoff to the bondholders is 0.5(100) + 0.5(70) = \$85 if the firm does not invest in the risky project and 0.5(100) + 0.5(5) = \$52.50 if the firm invests in the risky project. Thus, by investing in the risky project, the shareholders reduce the wealth of the bondholders by \$32.50. The shareholders themselves gain \$15, so that there is a net decline in total firm value of \$17.50. This is the aggregate loss due to moral hazard.

In this example, we assumed that the manager acted in the best interest of the shareholders. However, that is a questionable assumption too.[12] As an agent of the shareholders, the managers can do many things that may not be in the interest of the shareholders. For example, by inflating expenses, management can divert earnings from shareholders to management. Likewise, managers can discourage takeovers and thereby entrench themselves at the possible expense of shareholders. Managers may also select myopic and low-risk investment projects with a view toward protecting their positions and reputations.

You may have noticed that a critical assumption made in these examples is that the principal (the insurance company, the bondholders, or the shareholders) is unable to completely control the agent's behavior. If it were possible to costlessly observe the agent's actions, there would be no moral hazard. If the insurance company could precisely observe the insured, it would simply prohibit all actions detrimental to the car. It is because final outcomes do not unambiguously reveal the actions that may have influenced them that such proscriptions cannot be effectively written into contracts. Thus, for moral hazard to arise, it must be that: (i) the agent's actions (that affect the final outcome) cannot be costlessly observed by the principal, and (ii) there is some noise (exogenous uncertainty) that masks the agent's action in the final outcome.

Of course, the principal anticipates the agent's behavior. Thus, the principal attempts to design a contract that aligns the agent's incentives with her own. Deductibles and other coinsurance provisions in insurance contracts serve this purpose. Bondholders address moral hazard by limiting the firm's debt (the higher the debt/equity ratio, the greater is the inclination of shareholders to choose risky projects), by requiring collateral,[13] and by including in the debt contract covenants that restrict the borrower's actions. The interests of managers are aligned with the interests of shareholders through compensation contracts that include stock and stock options.

Another way to address moral hazard is to contract with the agent over extended time periods. Because of the possibility of reputational consideration, the agent may restrain self-interested behavior that is to the principal's detriment.[14] However, because lives are finite and because present consumption is usually preferred to future consumption, an agent's concern for reputation will not completely eliminate moral hazard.

It is important to understand that moral hazard is *not* the same as fraud. Most interesting cases of moral hazard do not involve illegal behavior. It is not illegal for shareholders to take on riskier projects than the bondholders would like. Nor is it illegal for a manager to invest in projects with faster paybacks than shareholders would like. Moral hazard may involve fraud, but it need not. It will almost always involve ethical considerations.

Agency and moral hazard issues, like asymmetric information, pervade much of this book. The chapters that make heaviest use of these ideas are Chapter 3 in which we discuss the role of banks and other financial intermediaries, Chapters 5 and 6 on spot lending issues, and Chapter 10 on deposit insurance.

12. See Jensen and Meckling (1976) and Mirrlees (1976). James Mirrlees, a British economist, was one of the pioneers in models of moral hazard in economics, and was awarded the Nobel Prize in Economics for his contributions.

13. Chan and Thakor (1987) and Boot, Thakor, and Udell (1991) show how moral hazard can be reduced by collateral. Stulz and Johnson (1985) examine the relationship between collateral and firm value.

14. See, for example, Diamond (1989), Holmstrom (1999), Hirshleifer and Thakor (1990), John and Nachman (1985), Song and Thakor (2006), and Thakor (2005).

Time Consistency

An issue that often crops up in moral hazard and adverse selection models is *time consistency*. To illustrate, suppose an employee expends effort to produce output on behalf of a principal. This output is affected by the agent's effort as well as some exogenous uncertainty that is beyond the agent's control. Thus, by observing the output the principal cannot be sure what effort the agent has taken. Suppose the principal is risk neutral and the agent is risk averse. Further, the principal must guarantee the agent some minimum level of expected utility[15] to ensure his participation. Finally, the agent would rather work less than more. The sequence of events is as follows: the principal gives the agent a wage contract, after which the agent expends some effort, following which the exogenous uncertainty is resolved, and then the output is realized. How should the agent be compensated?

If the principal could observe the agent's effort, the answer is simple. "Optimal" risk sharing is achieved if the principal pays the agent a fixed wage conditional upon the agent expending some prespecified effort, and nothing otherwise. This risk sharing scheme is optimal because it completely insulates the risk-averse manager from risk and imposes all of it on the risk-neutral principal. Because the effort is observable, it can be directly contracted upon. The agent will then do what the principal desires and receive a certain compensation that completely insures him against the randomness arising from the exogenous uncertainty.The principal receives the (random) output, but the randomness is costless because the principal is risk neutral.

If the agent's effort choice is unobservable, the above contract is unfeasible. The contract will be contingent on the only observable variable, the output. If the agent is promised a fixed wage, he avoids effort, so it is necessary to relate the wage to the output. This will motivate him to work harder to increase his share of the output. However, this approach to controlling the moral hazard has a cost. Since the agent is risk averse and his wage is uncertain, he will need to be compensated for the risk he bears. This will increase the principal's wage bill.

Now suppose that after the agent has expended his effort but before the output is realized, the principal has an opportunity to renegotiate the contract. Since the agent has already taken his effort, motivational concerns are irrelevant. The principal would be tempted to offer the agent a new wage that is fixed in amount (that is, independent of the output) and slightly less than the *expected* value of the agent's wage under the old contract. The risk-averse agent will gladly accept a slight reduction in his expected wage in order to rid himself of the income uncertainty inherent in the earlier contract. The risk-neutral principal is happy to save a little on her wage bill because the risk is a matter of indifference to her. Since both the

15. Because the agent is risk averse, it makes more sense to talk about a reservation expected utility than a reservation wage. To see this, suppose there was a choice between two wage contracts, W_1 and W_2, where W_1 pays \$144 for sure and W_2 pays \$400 with probability 0.5 and nothing with probability 0.5. Suppose we take \$121 as the minimum wage a risk-neutral agent will require to work for the principal. Then such an agent will accept either wage contract but will prefer W_2 since it has a higher expected value. On the other hand, a risk-averse agent with a utility function over wealth (W) given by $U(w) = \sqrt{W}$ will prefer W_1 to W_2. W_1 yields him an expected utility of $EU(W_1) = \sqrt{144} = 12$ utils, whereas W_2 yields him an expected utility of $EU(W_2) = 0.5\sqrt{400} = 10$ utils. If 11 utils is the minimum level of expected utility he needs in order to accept employment, then he will work only if he is offered W_1. You can see that we cannot define a minimum expected wage for such an agent since he does not evaluate his personal satisfaction by comparing expected wages. Rather, he computes the utility he expects to derive from each alternative.

principal and the agent are happy with this new arrangement, it's difficult to see why it would not replace the old one.

This is an example of a wage contract that is *time-inconsistent*. Although it seems like a good idea to negotiate a wage contract initially that conditions the agent's compensation on the realized output, such a contract will not work if both the agent and the principal recognize that they will subsequently want to renegotiate effect of the contract. The possibility of renegotiating the contract destroys the incentive effect of the contract. If the agent knows that his wage ultimately will be fixed, why should he work hard? To avoid this difficulty, it is necessary to build a *time-consistency* (or renegotiation-proofness) into the contract design. Contracts must be such that both parties to the contract should not have an incentive to renegotiate them.

To see how renegotiation-proofness affects contracts, consider the example of a bank-borrower relationship. The bank desires to protect itself against the borrower's incentive to increase the riskiness of the loan. It may use loan covenants that empower it to accelerate or call the loan if the borrower violates performance standards specified in loan covenants, often expressed in terms of financial ratios. The bank believes that this threat will induce the borrower to avoid excessive risk. However, when the bank is confronted with a violation of one or more of these covenants and threatens to accelerate the loan, the borrower offers a 50 basis point increase in the loan interest rate and offers assurances that the loan covenants will remain inviolate. The bank realizes that it can increase its reported profit by accepting the borrower's proposal and therefore withdraws its threat. To the extent that the borrower anticipates this behavior, the threat is not that the loan will be accelerated, but rather that the interest rate will be increased.

This is an example of a loan contract that is not renegotiation-proof. A renegotiation-proof loan contract would have specified interest rate penalties for minor loan covenant violations and would have included a loan acceleration provision only for violations so egregious (and informative) that the bank's best interest would call for the loan's termination regardless of possible enticements by the borrower.

Thus, contracts that are not renegotiation-proof are ultimately unsustainable. There is yet another aspect of time consistency that is unrelated to renegotiation-proofness. To illustrate, we shall use an *adverse selection* example. Suppose a bank is faced with two types of borrowers: good and bad. It can't distinguish between good and bad borrowers *a priori*, but if it could, it would lend only to the good borrowers. Suppose the borrower incurs a cost in applying for a bank loan. Moreover, the bank can discover whether a borrower who is good or bad by screening borrowers at some cost. If the bank does not screen, it charges a common interest rate to both types of borrowers and all borrowers who apply for credit. Borrowers know, however, that if the bank could distinguish among borrowers, it would lend to good borrowers exclusively. Now suppose the bank announces that it will screen all borrowers, so that it can sort out the bad borrowers and offer good borrowers a lower interest rate. Is this a time-consistent policy?

The answer is no. If borrowers believe that the bank will implement its policy, no bad borrower would apply for credit since the application cost would be wasted. However, they will anticipate this and infer that all applicants are good. But if they are all good, why incur a screening cost? Borrowers, in turn, anticipate this and realize there will be no screening, in which case *all* borrowers apply. But then it pays to screen! The result is an infinite regress and there is no equilibrium. We will have more to say about this issue in our discussion of credit rationing and bank regulation.

Nash Equilibrium

When agents transact with each other and each tries to selfishly maximize, they can be viewed as engaging in a *noncooperative game*. To describe the outcome, the concept of a *Nash equilibrium* has been proposed. Note first that by "equilibrium" we mean the attainment of some sort of a "steady state" in terms of the plans of action adopted by participants so that nobody can gain by unilaterally altering their plan of action. Before describing this equilibrium concept, notice that the outcome of the game depends on each player's actions. Moreover, each individual's actions will depend on what he thinks the adversary will do, since the final outcome is the collective resolution of individual actions. Thus, how each agent perceives the game will be played has an influence on each agent's choice of strategy and these choices determine the final outcome. To have an equilibrium, we cannot have erroneous beliefs. That is, if I take an action believing that you will do something, then you cannot do something else; if you do, the outcome cannot be an equilibrium. I would regret having made the decision and would wish to change it.

This intuitive notion is captured by the Nash equilibrium concept. Suppose there are n players engaged in a noncooperative game. Let S_i be the strategy (choice of action) of players i and let asterisks identify equilibrium strategies. Then the strategies $(S_1^*, S_2^*, \ldots, S_n^*)$ constitute a Nash equilibrium if, for every $i = 1, 2, \ldots, n$, S_i^* maximizes the personal welfare of agent i when all other agents play their equilibrium strategies. That is, suppose players 1 and 2 are engaged in a noncooperative game and strategies S_1^* and S_2^* represent a Nash equilibrium. Then, holding S_2^* fixed, player 1 cannot do better with any strategy other than S_1^*, and holding S_1^* fixed, player 2 cannot do better with any strategy other than S_2^*. We now illustrate this concept in the example below and in Figure 1.4.

Example 1.8 Suppose there are two prisoners who jointly committed a crime. There is insufficient evidence to convict either of them, unless one or both disclose information. The police, in an attempt to break their bond of silence, separately offer each the following deal. If prisoner 1 confesses and informs on prisoner 2 (who does not confess and inform on prisoner 1), then prisoner 1 will be freed. Let 4 represent the payoff equivalent to being set free after confessing. We assume that confessing and informing on his partner in crime causes the prisoner to feel a twinge of remorse, so that he enjoys 5 if he is freed without confessing. Of course, if prisoner 1 confesses and prisoner 2 does not, the latter will be convicted. Let 0 represent the payoff equivalent of being convicted. If both prisoners confess and inform, then both will be convicted, but the person who confesses and is still convicted receives a lighter sentence than one who remains silent and is convicted. Let 1 represent the payoff equivalent of being convicted despite confessing. Both prisoners know that if neither confesses, they'll both be set free. What will be the Nash equilibrium in this "prisoners' dilemma"?

Solution To answer this question let's first organize the payoffs to the various strategies in a matrix (known as the "strategic form" of this game). The first number in each cell is the payoff of prisoner 1 and the second number is the payoff of prisoner 2.

		Prisoner 2	
		Confess	Remain silent
Prisoner 1	Confess	1,1	4,0
	Remain silent	0,4	5,5

FIGURE 1.4 Strategic Form for Prisoners' Dilemma Game Prisoner 2

There are two Nash equilibria in this game: (i) both prisoners confess, and (ii) both players remain silent. To see why (i) is a Nash equilibrium, suppose prisoner 1 conjectures that prisoner 2 will confess. Then, if prisoner 1 confesses he gets 1, and if he remains silent he gets 0. So he confesses. On the other hand, suppose prisoner 2 conjectures that prisoner 1 will confess. Then, since his decision problem is same as that of prisoner 1, he too finds that confessing is optimal. Thus, (i) is a Nash equilibrium because, in choosing his strategy, each prisoner correctly conjectures the strategy of the other prisoner. Similarly, if each prisoner believes that the other will remain silent, then it is clearly best for each to remain silent. Thus, (ii) also is a Nash equilibrium.

Multiple Nash equilibria are common. Even though the two prisoners are clearly better off remaining silent, and even though they know this, it is possible for both to confess. This is because they can collude. Which equilibrium arises depends on trust among thieves.

The concept of Nash equilibrium is used extensively in the rest of this book. In particular, you will see quite a bit of it in Chapter 3, Chapter 5 and 6, and in the discussion of bank runs and deposit insurance in Chapter 10.

Revision of Beliefs and Bayes Rule

In this section we will discuss how a rational person would react to the arrival of new information. When a person does not know everything there is to know about something that will happen in the future, he can be viewed as formulating beliefs about what will happen. These beliefs can be described by a probability distribution. That is, as an incompletely informed person, you can say that you believe that there is some probability that outcome "a" will occur, some probability that outcome "b" will occur, and so on. Now, suppose some new information arrives. It does not inform you completely, but it adds to what you already know. The question is: how will you revise your original beliefs in the face of this new information? We illustrate this in the context of a specific example.

Example 1.9 Suppose you wish to determine the television channel on which you should watch the evening news to learn about the next day's weather. There are two main channels (say 1 and 2) that you can choose from. Your main criterion is the accuracy of the weather forecast, and you believe that the weather forecaster can be either "good" (g) or "bad" (b). Right now, you think that there is a 50-50 chance that the weather forecaster on either channel is good, that is, your (prior) belief is that the probability is 0.5 that the weather forecaster is g on either channel. You also realize that nobody is perfect, so that a good forecaster has a 0.8 chance of being right and a bad forecaster has a 0.5 chance of being right. Imagine for now that the forecasters on both channels give you "point estimates" (that is, they will tell you whether or not it will rain tomorrow) rather than probabilistic forecasts (for example, there's a 60% chance of rain). Suppose that the forecaster on channel 1 said last night that it would rain today and forecaster on channel 2 said that it would not. If you observe rain, how should you revise your beliefs?

Solution Clearly, it would not be wise to suddenly change your beliefs sharply and assert that the channel 1 forecaster is good and the channel 2 forecaster is bad. So, how should you proceed?

To answer this question, we need to formalize the belief revision process. *Bayes rule* is a statistical device that provides a formula to compute how beliefs should be revised. In essence, it tells us how a rational person should compute *conditional probabilities*. Suppose $x_1, \ldots, x_n$ are the possible realizations of the random variable x and $\Pr(x_i)$ is the prior (unconditional) probability that $x = x_i$, with x_i being some value chosen from $x_1, \ldots, x_n$. Similarly, y_i is some realization of the random variable y_i, which conveys information about x. Then, Bayes rule says that if you observe $y = y_i$, you should infer that the probability that $x = x_i$ is given by

$$\Pr(x_i|y_j) = \frac{\Pr(y_j|x_i)\Pr(x_i)}{\sum_{i=1}^{n}\Pr(y_j|x_i)\Pr(x_i)} \qquad [1.17]$$

The (unconditional) probability $\Pr(x_i)$ is known as a *prior belief* and the (conditional) probability $\Pr(x_i|y_j)$ is known as a *posterior belief*. In the context of our weather forecasting example, suppose we define

$$\Pr(\text{forecaster is good} \mid \text{he is correct}) \equiv \Pr(g \mid c)$$
$$\Pr(\text{forecaster is good} \mid \text{he is wrong}) \equiv \Pr(g \mid w)$$
$$\Pr(\text{forecaster is bad} \mid \text{he is correct}) \equiv \Pr(b \mid c)$$

and so on. Then,

$$\Pr(\text{channel 1 forecaster is good} \mid \text{he was correct in predicting rain})$$
$$= \Pr(g|c) = \frac{\Pr(c|g)\Pr(g)}{\Pr(c|g)\Pr(g) + \Pr(c|b)\Pr(b)}$$

$$= \frac{0.8 \times 0.5}{0.8 \times 0.5 + 0.5 \times 0.5} = 0.615.$$

Similarly,

$$\Pr(\text{channel 2 forecaster is good} \mid \text{he was wrong in predicting no rain})$$

$$= \Pr(g|w) = \frac{\Pr(w|g)\Pr(g)}{\Pr(w|g)\Pr(g) + \Pr(w|b)\Pr(b)}$$

$$= \frac{0.2 \times 0.5}{0.2 \times 0.5 + 0.5 \times 0.5}$$

$$= 0.286.$$

Thus, you now think that it is more than twice as likely that the channel 1 forecaster is good than it is that the channel 2 forecaster is good. Of course, you can wait until the next forecast and then see which (if either) of them is right. It is important to note that the posterior beliefs depend in a significant way on the prior beliefs. Thus, for example, if both forecasters predict rain tonight and it does rain tomorrow, you will *not* say that it is equally likely that they are good; you will still believe that there is a greater likelihood that the forecaster on channel 1 is good. We will see Bayes rule at work in Chapter 6.

References

Akerlof, George A., "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics*, August 1970, 488–500.

Arrow, Kenneth J., "The Role of Securities in the Optimal Allocation of Risk Bearing," *Review of Economics Studies*, 1964, 91–96.

Banker's Diary and Guide 1993, Warren Gorham Lamont, Division of Research Institute of America, Inc.

Bhattacharya, Sudipto, "Imperfect Information, Dividend Policy, and the 'Bird in the Hand' Fallacy," *Bell Journal of Economics and Management Science* 10, 1979, 259–270.

———, "Nondissipative Signaling Structures and Dividend Policy," *Quarterly Journal of Economics* 95, 1980, 1–24.

Black, Fisher, and Myron Scholes, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy* 81, 1973, 637–654.

Boot, Arnoud, Anjan Thakor, and Greg Udell, "Secured Lending and Default Risk: Equilibrium Analysis and Monetary Policy Implications," *The Economic Journal* 101–406, May 1991, 458–472.

Brealey, Richard, and Stewart Myers, *Principles of Corporate Finance*, Third Edition, McGraw Hill Publishing Co., 1988.

Brown, David P., and Robert H. Jennings, "On Technical Analysis," *Review of Financial Studies* 2, 1989, 527–551.

Chan, Yuk Shee, and Anjan Thakor, "Collateral and Competitive Equilibra with Moral Hazard and Private Information," *Journal of Finance* 42, June 1987, 345–363.

Debreu, Gerard, *Theory of Value*, Cowles Foundation Monograph 17, New Haven, Yale University Press, 1959.

Diamond, Douglas, "Reputation Acquisition in Debt Markets," *Journal of Political Economy* 97–4, August 1989, 828–862.

Fama, Eugene F., "Efficient Capital Markets: A Review of Theory and Empirical Work," *Journal of Business* 43, 1970, 383–417.

Fudenberg, Drew, and Jean Tirole, "Moral Hazard and Renegotiation in Agency Contracts," *Econometrica* 58–6, November 1990, 1279–1319.

Hirshleifer, David, and Anjan Thakor, "Managerial Reputation, Project Choice and Debt," *Review of Financial Studies* 5–3, 1992, 437–470.

Holmstrom, Bengt, "Managerial Incentive Problems–A Dynamic Perspective," *Review of Economic Studies*, January 1999.

Jensen, Michael, and William Meckling, "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure," *Journal of Financial Economics* 3, 1976, 305–360.

John, Kose, and David Nachman, "Investment Incentives, Risky Debt and Reputation in a Sequential Equilibrium," *Journal of Finance* 40, July 1985, 863–878.

Lehman, Bruce, "Fads, Martingales, and Efficiency," *Quarterly Journal of Economics*, February 1990, 1–28.

Markowitz, Harry, "Portfolio Selection: Efficient Diversification of Investments," Cowels Foundation Monograph 16, New Haven, Yale University Press, 1959.

Mirrlees, James, "The Optimal Structure of Incentives and Authority Within an Organization," *Bell Journal of Economics* 7–1, 1976, 105–131.

Narayanan, M.P., "Observability and Payback Criterion," *Journal of Business* 58, 1985a, 309–323.

———, "Managerial Incentives for Short-Term Results," *Journal of Finance* 40, 1985b, 1469–1484.

Ross, Stephen A., "On the Economic Theory of Agency: The Principle of Similarity," Proceedings of the NERB-NSF Conference on Decision Making and Uncertainty, 1974.

Samuelson, Paul, "Risk and Uncertainty: A Fallacy of Large Numbers," *Scientia* 57, 1963, 1–6.

Song, Fenghua, and Anjan V. Thakor, "Information Control, Career Concerns and Corporate Governance," *Journal of Finance* 61–4, August 2006, 1845–1896.

Spence, A. Michael, "Job Market Signalling," *Quarterly Journal of Economics*, 1973, 355–374.

———, "Competitive and Optimal Responses to Signals: An Analysis of Efficiency and Distribution," *Journal of Economic Theory* 7, 1974, 296–332.

Stulz, Rene Herb Johnson, "An Analysis of Secured Debt," *Journal of Financial Economics* 14–4, December 1985, 501–522.

Thakor, Anjan V., "Do Loan Commitments Cause Overlending?" *Journal of Money, Credit and Banking* 37–6, December 2005, 1067–1100.

PART ♦ II

What Is Financial Intermediation?

This page intentionally left blank

CHAPTER • 2

The Nature and Variety of Financial Intermediation

"Don't it always seem to go that you don't know what you've got 'til it's gone?"
Joni Mitchell

Glossary of Terms

Euro: Common currency adopted by many member countries of the European Union.

Yield Curve: Relationship between yield to maturity and maturity on debt instruments identical in all respects except maturities (see Chapter 4).

Duration: A measure of how long an investor must wait to receive payment on a bond. For bonds that repay only principal (zero coupon bonds), duration equals maturity. For coupon-paying bonds, duration is always shorter than maturity (see Chapter 4).

Spot Rate: The current yield to maturity on a bond of a given maturity (see Chapter 4).

Liquidity Premium: The amount by which the yield on a bond must be grossed up to compensate investors for their inability to convert the bond into cash at a moment's notice and without loss relative to the bond's true value (see Chapter 4).

Consumer Loans: Loans made to individuals and families. These are primarily installment loans.

Commercial Loans: Loans made to corporations. Often referred to as Commercial and Industrial (C&I) loans.

Contingent Claims: Claims that may be made in the future, contingent on the realizations of some states.

Federal Funds: Funds in the interbank loan market. When a bank "sells" federal funds, it is lending (usually on an overnight basis) to another bank an amount that covers a part or all of that bank's shortfall in reserves; banks are required to keep a certain fraction of their deposits as liquid reserves.

Surplus: Proceeds from the sale of equity and securities in excess of their par value, plus earnings retained until the surplus account equals the common stock account.

Cash and Due: Coin and currency in the bank's vaults, reserves on deposit with the Federal Reserve and with other banks, and checks deposited by customers on which funds have not yet been collected from the paying bank.

Allowance for Loan Losses: An allowance made to absorb anticipated (expected) future loan losses. The amount allocated for loan losses is part of the bank's net worth. An allowance for loan losses is a charge against current income and it increases the bank's *loan loss reserve*. Writeoffs of existing loans reduce the bank's loan loss reserve.

Undivided Profits and Reserves: Part of the bank's net worth.

Types of Life Insurance Policies: There are basically four types of life insurance policies—ordinary, industrial, group, and credit. An *ordinary life insurance* policy is the kind most individuals have. It involves monthly, quarterly, semi-annual, or annual premium payments and specified benefit payment upon death. *Industrial insurance* comes in small denominations with weekly or monthly premiums usually collected at the insured's home by an agent. *Group insurance* covers a number of people—employees of a particular firm or union members, for example—under a single policy issued without medical examination. *Credit life insurance* is individual or group term insurance that provides for repayment of the insured's debt in the event of the insured's death.

Introduction

This chapter focuses on the variety of services provided by financial intermediaries. Banks are members of an expansive industry that provides a dazzling variety of financial services. The broader financial services industry includes institutions as different as commercial banks, savings institutions, and credit unions, all of which finance their assets with deposits, and government agencies, pension funds, loan sharks, pawnbrokers, lotteries, insurance companies, mutual funds, hedge funds, and private-equity pools. To this list we could add organized exchanges for trading stocks, futures, options, bonds and commodities, pari-mutuel betting institutions, credit-rating agencies, and the list can be extended almost effortlessly.

What all these *financial institutions* have in common is the processing of risk and its subtle complement, information. Financial intermediaries produce information for two kinds of applications: (i) to match transactors like a marriage broker would, and (ii) to manage risks and transform the nature of claims as when a bank produces credit information to control a borrower's credit risk. In producing information for application (i), the intermediary acts as a *broker*, whereas in producing information for application (ii), it acts as a *qualitative asset transformer*.

Our plan in the rest of this chapter is as follows. First we define financial intermediaries (F.I.s) and discuss brokerage and asset transformation services. We also provide a list of the different types of services that intermediaries provide in each of these two basic groups. Next we provide some key statistics about financial intermediaries. Then we discuss the main types of depository intermediaries: commercial banks, thrifts (savings and loan associations and mutual savings banks), credit unions, and venture capitalists. The next section discusses nondepository financial intermediaries: finance companies, insurance companies, pension funds, mutual funds, and investment banks. We cover the role of the government next and then turn to "peripheral" financial intermediaries, including pawnbrokers and loan sharks. Appendix 2.1 discusses valuation problems related to F.I. balance sheets, and Appendix 2.2 provides a summary of key regulations affecting banks.

What Are Financial Intermediaries?

Definition: As the name suggests, financial intermediaries (F.I.s) are entities that intermediate between providers and users of financial capital. F.I.s are typically multifaceted, and their activities therefore can be understood from a variety of vantage points. For example, in contrast with nonfinancial firms, F.I.s hold relatively large quantities of financial claims as assets. Thus, whereas the manufacturing firm holds inventories, machines, and patents as assets, the F.I. holds contracts of the indebtedness of their clients as assets. Both finance their assets by selling their own debt and equity; there is no compelling distinction between F.I.s and others on the right-hand side of the balance sheet, except that F.I.s tend to be more leveraged. Here we have a balance sheet perspective on the uniqueness of financial intermediation. Whereas both F.I.s and other types of business finance assets with debt and equity, F.I.s tend to hold financial claims as assets whereas others are more committed to physical assets. In Appendix 2.1, we provide a further discussion of the balance sheets of F.I.s.

Why Do We Have F.I.s?: This is tantamount to asking: What do F.I.s do that could not be done without them? The answer to this for *any* firm, financial or nonfinancial, is found in the flow of goods and/or services produced by the firm. After all, a firm not only selects its assets and liabilities but also manages them so as to assure the realization of the potential cash flows. That is, the (nonhuman) assets appearing on the balance sheet are combined with various kinds of labor inputs to produce the cash flows conventionally attributed to the assets. The manufacturer reshapes, transforms, and transports various raw materials and semifinished goods into more highly refined and more advantageously located goods. The services of machines and processes recorded on the balance sheet are combined with labor services to produce inventory of more highly refined goods.

What is the analog for the F.I.? How does it combine its resources to produce financial services? A facile answer is that F.I.s borrow on the one hand and lend on the other. But this answer is incomplete because it doesn't explain why we need F.I.s to bring borrowers and lenders together. That is, if I wanted to borrow some money, why don't I simply put an ad in the newspaper and invite people to lend to me at interest rates that I could negotiate with them? While this may seem to some like a foolish thing to do, the key is to understand why it isn't (normally) done, rather than to dismiss it outright. After all, is it that different from a homeowner putting up his house "for sale by owner," rather than through a real estate agent? Why is the selling of a house different from the selling of one's indebtedness (borrowing money)? Even in countries where there is not (explicit) deposit insurance, people deposit money in banks, which in turn lend this money to people like you and me. So, why aren't those depositors willing to transact directly with prospective borrowers?

The key to understanding this issue is that we live in a world of imperfect information. People would rather deposit their money in a bank than lend it directly to a stranger because they feel they "know" the bank better. It is this line of reasoning that we wish to explore further, with emphasis on the information-based financial services produced by a F.I. In borrowing and lending, the F.I is joining unfamiliar, but well-suited and complementary transactors, much like the marriage broker would. The F.I. also is allocating credit presumably to its highest and best uses while reconfiguring the attributes of the financial claims held by its clienteles.[1] These activities are so fundamental to financial intermediation that they are accorded special labels, the former being referred to as "brokerage" whereas the latter is called "qualitative asset transformation" (QAT). Let us explain each in turn.

The Brokerage Function of F.I.s: Brokerage activities of F.I.s involve the bringing together of transactors in financial claims with complementary needs. The broker is usually compensated with a fee for performing this service. The broker's stock-in-trade is information, and its special edge in performing this service derives from special skills in interpreting subtle (that is, not readily observable) signals, and also from the reusability of information. That is, a broker has two advantages as an information processor. First, it possesses/develops special skills in interpreting subtle (not readily observable) signals. Second, it takes advantage of cross-sectional (across customers) and intertemporal (through time) information reusability. For example, a real estate broker typically has better information than the average home buyer or seller about supply and demand conditions in a given market and is able to reuse this information on many transactions.

For the broker, the matching of buyers and sellers does *not* involve the broker as a principal in the purchase (sale). Thus the used car dealership typically goes beyond the broker's role in that it will *purchase* used autos for resale. If it merely identified potential buyers (sellers) for counterparties, it would then be a broker. Likewise, the marriage broker fits our description of a broker, but the typical stockbroker does not. Once a broker serves as principal and buys (sells) the asset for eventual resale (repurchase), it accepts the risk that the market may reprice the asset, and it therefore transcends the more limited role of the matchmaker.

The broker helps resolve informational problems that exist before the two sides to the transaction enter into a contract, i.e., the broker helps resolve *precontract informational asymmetry*. Moreover, the broker also helps resolve informational problems that may arise after the contract is entered into, i.e., the broker helps resolve *postcontract informational asymmetry*.

1. F.I.s also engage in clearing and storage activities that are still more closely analogous to manufacturing. These asset "servicing" activities include collecting, tracking, and remitting payments on mortgages, consumer credit, and other claims, as well as traditional safekeeping.

Precontract Informational Asymmetry and Brokerage: Precontract information asymmetry involves two kinds of information problems: adverse selection and duplicated screening. We will discuss each in turn.

Adverse Selection and Brokerage: In transactions involving F.I.s, adverse selection problems abound. For example, a borrower will wish to overstate his credit worthiness to potential lenders in order to make himself look like a low-credit-risk borrower. And if the lender raises the loan interest rate in order to be compensated for the higher credit risk associated with borrowers, who misrepresent their creditworthiness, the borrowers most likely to drop out are the *low*-credit-risk borrowers who may either have better credit alternatives or be simply unwilling to borrow at the higher interest rate. Consequently, the lender is left with only the high-credit-risk borrowers.

An F.I. like a bank can help deal with this adverse selection problem by performing the brokerage function of credit analysis to sort out borrowers of different credit risks. That is, in this case the broker specializes in credit analysis or develops the skills to process/interpret various types of credit information. This allows it to intermediate between borrowers and lenders and minimize adverse selection problems.

Duplicated Screening, Information Reusability and Brokerage: Duplicated screening refers to situations in which individuals can resolve adverse selection at a cost, but there is wasteful expenditure of costly screening resources because multiple individuals end up doing the same screening. An F.I. can help avoid such duplication by exploiting the power of information reusability. This can be illustrated through the example given below.

Consider 100 men and 100 women searching for the "perfect" marriage partner. In order to become fully informed, each woman will need to evaluate each of the 100 men, and likewise for each of the men. Now suppose that each such evaluation (sampling) results in a fixed cost of say, \$25. Then the total cost for all participants to become fully informed would be \$500,000 (that is, $2(100 \times 100 \times 25)$). Or, if we let x represent the size of the side of a square grid (100 people in this example), and c the fixed sampling cost per unit (\$25 in this example), we have the result that the total cost equals $2cx^2$.

Now enter the broker! To establish a level playing field and suppress consideration of the broker's special skills, we assume the evaluation cost per unit remains unchanged at \$25. However, the broker will need to examine each of the participants only once and hence its total cost of becoming informed is $2cx$, or \$5,000. Assuming the information is distributed at negligible cost, the saving due to the introduction of the broker is approximated by

$$S = 2cx(x - 1),$$

or \$495,000 in the example. To be sure, the broker will expect to earn a profit, but this cost is redistributive rather than dissipative (resource consuming), and potential competition can be expected to limit the profit in any case. Thus the saving associated with having a broker increases exponentially (the square) with the size of the grid, and linearly with the sampling cost per unit. At the margin ($dS/dx = 2c[2x - 1]$), the saving is increasing as the size of the grid expands.

The savings, due to the broker, derive from a peculiarity of information: its use does not result in its consumption. Most goods and services are transformed into waste as a result of being used. This is not true with information, and this idiosyncracy is the key to understanding the broker's role. If the marriage broker composes a report on a particular candidate, I can use the information without in any way compromising your ability to use the same information. The same is true for a report written by a security analyst, or for a telephone book. This extraordinary reusability of information is what makes it compelling to have a broker, and the larger the grid, the greater the potential saving associated with reusing information.

In this discussion, we did not assign the broker any special advantage or skill relative to the lay person in information evaluation. If such a relative advantage exists, then let C_b = broker's evaluation cost and C_o = others' evaluation cost, with $C_o > C_b$. Then the saving due to the broker is $S = 2x[C_o x - C_b]$, with $C_o > C_b$, and the saving attendant to using the broker grows with the gap, $C_o - C_b$. That is, higher information processing skill accentuates the broker's relative advantage.

Some Further Thoughts on the Power of Information Reusability and the Value of Brokerage: To cement our understanding of the power of information reusability, consider one more example. Think of a very large geographic grid in which each intersection represents a potential oil well. Now suppose there are many oil drilling entrepreneurs, and further suppose that after drilling a dry hole the law requires that the landscape be restored to its initial condition. Thus, there is no way to know if a particular location has been drilled unless there is an operating well at a particular location. If a broker simply collects and disseminates information about the drilling activities of each explorer, the cost of redrilling dry holes can be eliminated. Without the broker, society will bear the unnecessary cost of searching for oil in locations known to be unproductive. This aspect of information is called *cross-sectional reusability*; the same information can be utilized across a number of different users. Information reusability also has an *intertemporal* aspect; it can be reused through time. For example, a bank that learns something about a borrower while processing its first loan application can use at least some of that information in processing future credit requests from the *same* borrower.

A second aspect of brokerage relates to the *observability* of objects of search. When the object of search is trivially observable, as in the case of a person's telephone number or the address of a dry hole, the skills of the broker are of little importance. But let us be a little more precise in explaining what we mean by "trivially observable." Think of the problem of retaining an expert to assist you in the purchase of thoroughbred horses. Suppose that you are particularly interested in three traits of candidate horses—their racing records, conformation, and blood lines. Now imagine there are numerous experts available and suppose we ask each to report on the three traits of a sample horse. We then create a frequency distribution for each trait. What would we expect to observe among these frequency distributions? Because the racing records are well-defined and a matter of public record, deviations around the mean should be negligible. Observers will not dispute how many times a particular race horse has come in first, second, and so on, no more than they would dispute its age, weight, or height.

However, breeding and conformation are a very different kettle of fish. With regard to these attributes, we would expect each agent to report a different description of the subject horse. Since the ideal against which conformation is judged is multidimensional and somewhat loosely defined, each observer's characterization will be distinctive and the consequent frequency distribution will have considerable variance. Likewise for bloodlines. The facts relating to forebears may be indisputable, but the value of particular forebears is judgmental; the choice among observers thus becomes important.

It is the subtlety, vagueness, or cost of observing the objects of search that elevates the importance of broker skills. To the extent that the objects of search are trivially observable, we should wish to employ less astute observers. If all observers produce the same description, clearly we should reserve the most astute brokers for those searches where judgments matter.

The observability issue helps us to understand the striking hierarchy of brokers in society, ranging from phone books at one extreme to marriage brokers and investment bankers at the other. Indeed, investment bankers and marriage brokers have a good deal in common in that they both address the pairing of transactors on the basis of subtle attributes. If the investment banker were limited to *pro forma* financial statements and projecting cash flows, its role and compensation would both be diminished. But presumably the investment banker addresses more complicated issues of compatibility based on corporate cultures, strategic intent, succession, operating synergies, and similar nuances. Even the placing of securities requires a knowledge of buyers and sellers and how they view counterparties as well as the many details of securities' attributes, such as sinking fund provisions, collateral, and stochastic duration considerations. This explains why the reputation of the investment banker is critically important, whereas the publisher of the Yellow Pages is virtually anonymous.

To summarize, for a given attribute, the larger the grid, the more compelling the need for the broker. For a given size grid, the less readily observable the object of search, the more important the skills and reputation of the broker.

An important aspect of brokerage is that it can be performed without processing substantial risk. Information can be purchased for resale without exposing the broker in the way QAT (qualitative asset transformation) does. To be sure, if the broker produces information before it is sold, demand uncertainty can result in losses. But information can be presold, at least in principle. The broker also exposes its reputation whenever falsifiable representations are made in connection with its sale of information. But the risk is material only to the extent that objects of search are observable with difficulty. In principle then, brokerage services can be produced risklessly, and in any case the processing of risk is not central to the production of brokerage services. This is not the case with QAT.

Postcontract Informational Asymmetry and Brokerage: In many transactions, one party to the transaction can take actions during the course of the contractual interaction that damage the interest of the other party. The reason why such behavior is possible is that these actions are "hidden" from the injured party and cannot be directly controlled or prevented. Such informational asymmetry is associated with moral hazard, discussed in Chapter 1.

Moral hazard is quite prevalent. It is encountered in insurance, where the insured may underinvest in costly efforts to prevent adverse outcomes because the insurer absorbs the resulting loss. It is encountered in banking, where borrowers may choose excessively risky projects because the bank bears a disproportionate share of the downside risk.

The F.I.'s special skills in monitoring attenuate moral hazard. For example, banks monitor their borrowers by periodically examining the borrower's business and its financial condition and intervening in operating strategy when necessary. Insurance companies design insurance contracts and use ex post pricing adjustments to deter moral hazard. Venture capitalists use the threat of transfer of control to ensure that the entrepreneur's incentives do not stray too far from investors' desires.

Thus, moral hazard provides a powerful source of economic value for the F.I. to emerge as a broker that can help diminish the losses due to moral hazard.

Figure 2.1 below summarizes the different informational problems that create a role for the broker.

Qualitative Asset Transformation: Think of a world without intermediaries, as we did at the beginning of this discussion about the role of intermediaries. Suppose some individual wishes to borrow for the purpose of purchasing a house. The borrower must find a counterparty willing to hold a mortgage, which is a claim with a number of less desirable attributes. For example, there is no active secondary market in individual mortgages, with resultant illiquidity, and wide bid-ask spreads. The mortgage typically comes in large and irregular unit sizes. It typically has a long and uncertain duration, which is to say it may remain unredeemed for 30 years, but it can be repaid at virtually any time at the borrower's discretion and typically without prepayment penalty. Moreover, the mortgage carries with it default risk, and in the event of default, managing the collateral can be expensive. In all, the mortgage is a homely claim.

Enter the F.I.! It purchases the mortgage and finances the purchase with the issuance of a liability called a deposit. The deposit, in contrast to the mortgage, is almost infinitely divisible, highly liquid, and has little default risk. The F.I. effectively swaps deposits for mortgages, thereby modifying the claims held by its clientele. The F.I. is rewarded for this service with interest rate spread between deposits and mortgages.

Among the asset attributes most commonly transformed by F.I.s are duration (or term-to-maturity), divisibility (or unit size), liquidity, credit risk, and sometimes numeraire (currency identity). Typically, the intermediary will shorten the duration of the claims of its clients by holding assets of longer duration than its own liabilities; it will reduce the unit size of the claims of its clients by holding assets of larger unit size than its liabilities; it will enhance the liquidity of the claims of its clients by holding assets that are more illiquid than its liabilities; and it will reduce credit risk by holding assets that are more likely to default than its liabilities. By holding assets denominated in a currency other than its liabilities, it alters the numeraire of the assets of its clients.

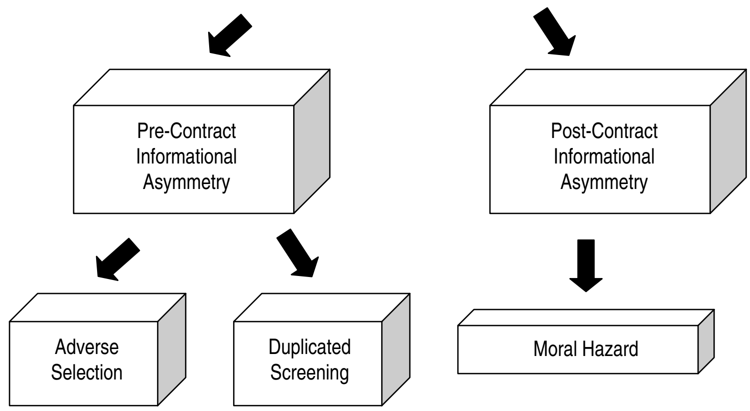


FIGURE 2.1 Key Information Problems Addressed by F.I.s

QAT and Risk: But notice that every such asset transformation performed by the F.I. requires a *mismatch* with regard to that attribute on the F.I.'s balance sheet. For example, if the duration of the F.I.'s assets and liabilities are perfectly matched, it cannot have altered the duration of the assets of its clients. Only by absorbing the longer duration assets in exchange for shorter term liabilities can the F.I. reduce the duration of claims held by its customers. This is important because the mismatch on the F.I.'s balance sheet reflects an acceptance of some type of risk, at least initially, by the F.I.

If the F. I. holds Euro-denominated assets and U.S. dollar–denominated liabilities, it will be exposed to variations in the dollar/Euro exchange rate. If it holds long-term assets financed with short-term liabilities, it will be exposed to interest rate risk, whereby changes in the shape and position of the yield curve will affect the F.I.'s cash flows. Even changing the unit size of claims cannot be done without a mismatch and a consequent acceptance of risk. If the unit size of assets is larger than that of liabilities, the purchase and sale of corresponding claims cannot be perfectly synchronized and hence the F.I. accepts a form of inventory risk.

The case of duration transformation is particularly instructive. The yield curve is thought to be a ''biased predictor'' of future spot interest rates owing to a (liquidity) premium attached to long-duration claims. That is, borrowers typically prefer to borrow long term and lenders typically prefer to lend short term. This theory of the term structure of interest rates is usually associated with Sir John Hicks, a British Nobel Laureate economist. But if we introduce F.I.s into such a world and assume that they are indifferent to the duration of a claim, they would be able to finance the purchase of long-term assets with short-term liabilities and profit from doing so. Indeed, absent other impediments, intermediaries would continue to perform this transformation until the liquidity premium is bid down to the marginal cost of intermediating. The existence of this form of asset transformation supports the Hicksian view of the yield curve. Without a liquidity premium at the outset, there would be no incentive for the F.I. to perform duration transformation. If the yield curve was an unbiased predictor of future spot interest rates, there would be no profit in performing duration transformation.

Whatever the form of the QAT, a mismatched balance sheet is implied, and this in turn implies the acceptance of some form of exposure. This is the sense in which risk is integral to QAT. In managing this risk, there are basically three alternatives available to the F.I. It can diversify the risk, it can shift the risk to others, or it can passively accept the exposure. The shifting of risk to others involves the use of claims such as swaps, forward contracts, futures and options, and in principle, but rarely in practice, all of the exposure associated with the QAT can be transferred to others with the appropriate risk-shifting instruments. However, in this case the QAT reverts to brokerage. The F.I. has merely transferred risk among its clients, no matter how convoluted the transactions. In the case where the risk is diversifiable, presumably the F.I. performs this diversification on behalf of clients whose wealth is too small relative to the unit size of claims to diversify on their own. It is widely believed that this is a major rationale for mutual funds.

Although we distinguish between brokerage and asset transformation as distinct types of intermediation services, the truth is that both are performed by the same intermediaries and sometimes in combination. Take for example a duration-transforming F.I. that finds it is too mismatched for comfort and consequently proceeds to lengthen the duration of liabilities while simultaneously shortening the duration of its assets. In fact, it is changing the mix of its activities from more to less

QAT and from less to more brokerage. In the limit, if the F.I. achieves a perfect duration match of its assets and liabilities, it will have become a pure broker.

Or consider an investment banker with two types of underwriting contracts, the "firm commitment" contract and the "best efforts" contract. The form involves the banker purchasing a firm's securities for resale. This is clearly a QAT contract. The banker provides the issuing firm with a *prix fixé* before the public has committed to purchase the securities. By contrast, the best efforts contract merely commits the bankers to make an honest effort to sell the securities for the best realizable price, without any further assurances. The best efforts contract commits the banker to provide brokerage services, and the banker will typically receive a fee without accepting any exposure relating to the price of the securities. Figure 2.2 lists the various services provided by F.I.s under brokerage and QAT. This list is suggestive, not exhaustive.

The Variety of Financial Intermediaries

There are many ways to classify the many different types of F.I.s. In the previous section we classified them based on the nature of the services they provide. We can also classify them based on whether or not they finance their activities with deposits.

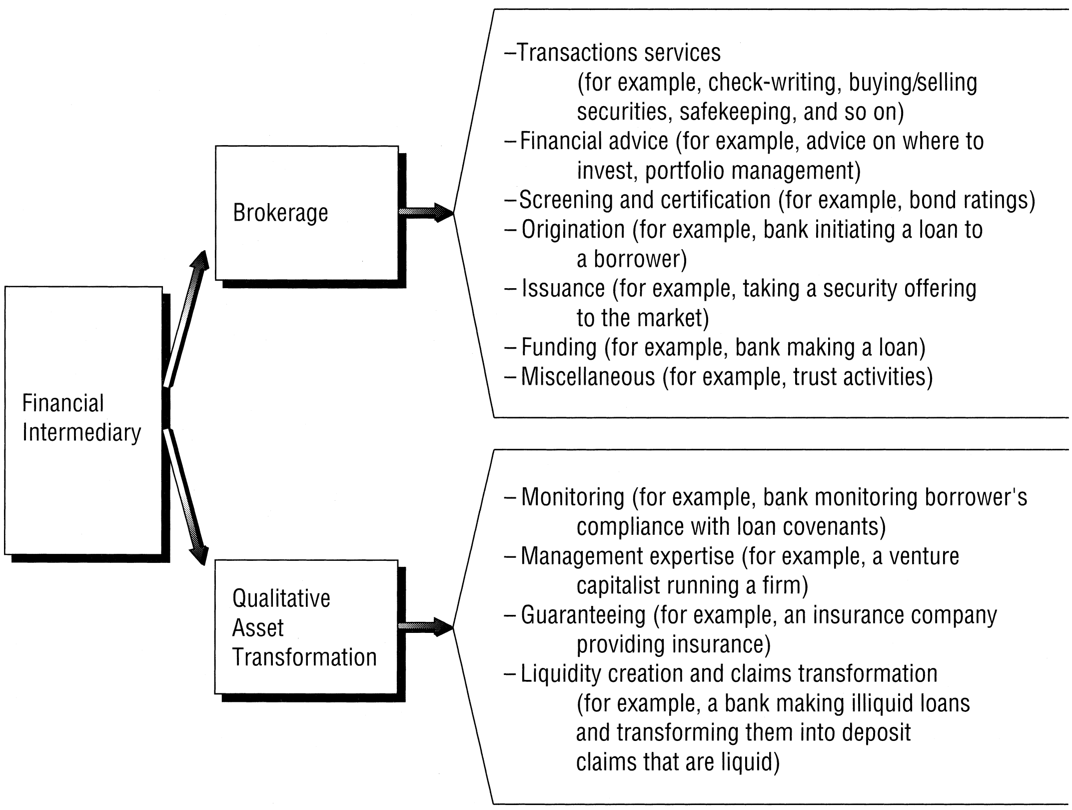


FIGURE 2.2 Services Provided by Financial Intermediaries

F.I.s that finance (at least partly) with deposits are called deposit-type or depository F.I.s, whereas those that do not finance with deposits are called nondepository F.I.s. Jointly, depository and nondepository F.I.s have at their command an enormous volume of assets. Table 2.1 lists the assets of the various types of F.I.s, and also depicts their growth during 1980–2003. It is noteworthy that the assets of *all* types of F.I.s, except savings institutions, have exhibited striking growth.

The distinctions between banks and other depository F.I.s and between depository and nondepository F.I.s have become blurred during the last two decades. The distinctions between investment banks and commercial banks have diminished as the latter have responded to competition from the capital market by increasing loan sales, providing financial guarantees, and directly placing securities for customers. The distinctions between depository and nondepository institutions have also become blurred as the latter have increasingly offered products and services that compete with those of commercial banks. Consequently, individuals are increasingly

TABLE 2.1 Total Assets of Financial Intermediaries at Year-End

Panel A: Total Assets Expressed in Billions of Dollars

Financial Intermediary	1980	1985	1990	1995	2000	2003
Commercial Banks	$1,704	$2,484	$3,337	$4,494	$6,469	$7,812
Savings Institutions	792	1,287	1,323	1,013	1,218	1,475
Life Insurance Companies	479	826	1,351	2,064	3,136	3,823
Private Pension Funds	470	848	1,627	2,889	4,423	4,194
State and Local Pension Funds	198	405	801	1,303	2,290	2,284
Finance Companies	202	352	574	672	1,140	1,381
Money Market Funds	76	244	493	741	1,812	2,016
Mutual Funds	58	252	608	1,853	4,435	4,665
Credit Unions	69	137	217	311	441	617
Financial Intermediaries' Total Assets	**$4,048**	**$6,835**	**$10,331**	**$15,340**	**$25,364**	**$28,267**

Panel B: Total Assets Expressed as a Fraction of Financial Intermediaries' Total Assets

Financial Intermediary	1980	1985	1990	1995	2000	2003
Commercial Banks	0.42	0.36	0.32	0.29	0.26	0.28
Savings Institutions	0.20	0.19	0.13	0.07	0.05	0.05
Life Insurance Companies	0.12	0.12	0.13	0.13	0.12	0.14
Private Pension Funds	0.12	0.12	0.16	0.19	0.17	0.15
State and Local Pension Funds	0.05	0.06	0.08	0.08	0.09	0.08
Finance Companies	0.05	0.05	0.06	0.04	0.04	0.05
Money Market Funds	0.02	0.04	0.05	0.05	0.07	0.07
Mutual Funds	0.01	0.04	0.06	0.12	0.17	0.17
Credit Unions	0.02	0.02	0.02	0.02	0.02	0.02
Financial Intermediaries' Total Assets	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**

Source: U.S. Census Bureau, Statistical Abstract of the United States: 2004–2005.

turning to mutual funds rather than bank deposits for transactions and investment purposes. These developments can be seen in the data provided in Tables 2.2 and 2.3.

The shifting market shares of various institutions in the consumer loan market are reflected in the data provided in Table 2.4. Commercial banks are still the biggest players in the consumer loan market. The ten largest commercial banks in consumer lending are shown in Table 2.5. The share of different financial institutions in total credit is shown in Figure 2.3. Having provided you with a glimpse of the market shares and sizes of the various types of institutions, we now move on to a description of each of these institutions in the next section.

TABLE 2.2 Various Mutual Fund Statistics (In Billions of Dollars or in Percentage)

	1980[1]	1990	2000	2004
Dollars Invested in Mutual Funds	$134.8	$1,065.2	$6,964.7	$8,106.9
Mutual Funds Share of I.R.A. Market[1]	14.0%	21.8%	46.7%	42.8%
Penetration of Mutual Funds Among U.S. Households	5.7%	25.0%	49.0%	48.1%

[1] Mutual funds share is from the mid-1980s.
Source: Investment Company Institute 2005 Fact Book.

TABLE 2.3 U.S. Mutual Fund Industry Total Net Assets (In Billions of Dollars)

	1980[1]	1990	2000	2004
Long-Term Funds				
Equity Funds	$44.4	$239.5	$3,961.9	$4,384.1
Hybrid Funds		$36.1	$346.3	$519.3
Bond Funds	$14.0	$291.3	$811.2	$1,290.3
Money-Market Funds	$76.4	$498.3	$1,845.3	$1,913.2
Total Net Assets	$134.8	$1,065.2	$6,964.7	$8,106.9
Number of Funds	564	3,079	8,155	8,044

[1] All funds were reclassified in 1984 and a separate category was created for hybrid funds.
Source: Investment Company Institute 2005 Fact Book.

TABLE 2.4 Market Share of Consumer Loans (In Percentage)

	1–4 Family Mortgages			*Consumer Credit*		
	1990	2000	2004	1990	2000	2004
Commercial Banks	16.5	18.9	19.4	46.3	32.0	33.2
Savings institutions	23.0	11.6	10.8	6.0	3.7	4.3
Life Insurance Companies	0.5	0.1	0.1	–	–	–
Finance Companies	1.5	2.5	2.4	16.8	12.7	17.2

Source: Federal Reserve Statistical Release: Flow of Funds Accounts of the U.S. 1985–1994 and 1995–2004.

TABLE 2.5 Top Ten Banks Based on Total Assets in November 2005

Name	City, State	Total Assets (Billions of Dollars)	Total Deposits (Billions of Dollars)
1. Citigroup Inc.	New York, NY	1,472.8	581.1
2. Bank of America Corporation[1]	Charlotte, NC	1,314.9	655.7
3. JPMorgan Chase & Co.	New York, NY	1,203.0	535.1
4. Wachovia Corporation[2]	Charlotte, NC	549.4	322.8
5. Wells Fargo & Company	San Francisco, CA	453.5	289.0
6. USBC North America Holdings Inc.[*]	Prospect Heights, IL	372.6	114.1
7. U.S. Bancorp	Minneapolis, MN	206.9	120.8
8. Sun Trust Banks, Inc.	Atlanta, GA	172.4	113.7
9. Citizens Financial Group, Inc.[*]	Providence, RI	148.5	97.3
10. National City Corporation	Cleveland, OH	146.6	83.4

[1] Reflects Bank of America Corporation's pending acquisition of MBNA Corporation.
[2] Reflects Wachovia Corporation's pending acquisition of Westcorp.
[*] Financial information as of June 30.
Source: SNL Financial.

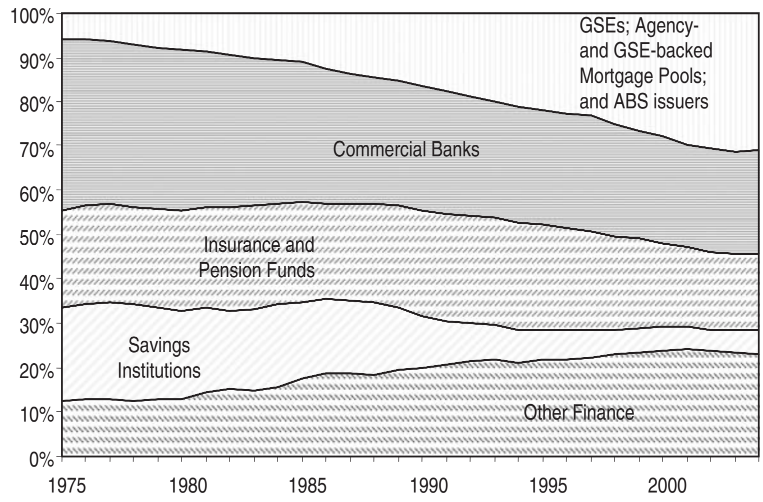


FIGURE 2.3 Share of Financial Institutions in Total Credit

Depository Financial Intermediaries

Depository institutions operate with high leverage, so that even a small return on total assets translates into a high return of equity. Figure 2.4 graphs the behavior through time of bank equity capital as a percentage of total assets. The figure illustrates the post-World War II upward drift in the net-worth-to-total-assets ratio through the 1960s then the long-run decline in the net-worth-to-total-asset ratio of banks until about 1980, followed by an increase in this ratio thereafter, for reasons that will be discussed later. In Figure 2.5 we provide information on the return on assets and the return on equity at commercial banks. This figure highlights the effects leverage has on the translation from a return on assets to a return on equity. For instance, in 2004 return on assets for commercial banks was about 1.31 percent, whereas return on equity was 13.82 percent.

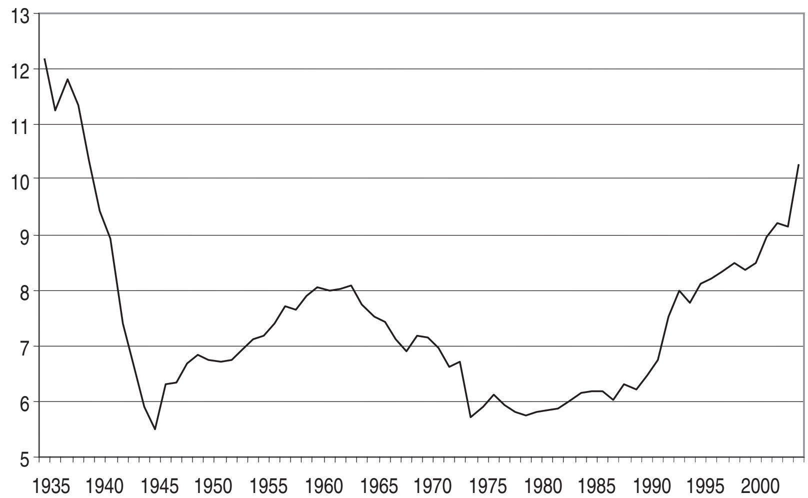


FIGURE 2.4 Bank Equity Capital as a Percent of Total Assets
Source: FDIC Quarterly Banking Profile.

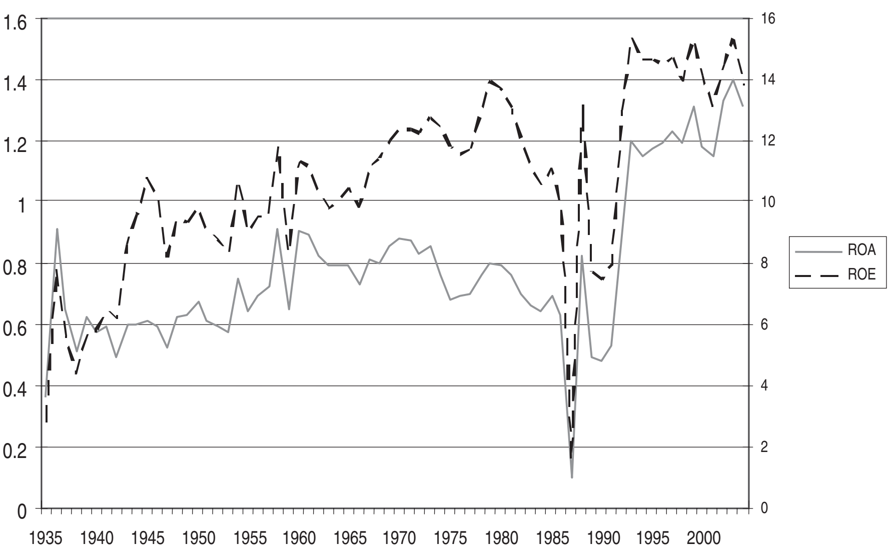


FIGURE 2.5 Commercial Bank Profitability
Source: FDIC Quarterly Banking Profile.

Commercial Banks

Commercial banks are widely considered the center of the financial intermediation universe because of their role in administering the community's payments, and also because commercial banks are used to transmit monetary policy impulses originating with the central bank.[2] Their sheer size and ubiquity provide yet another basis for according commercial banks special attention.

Most commercial banks operate with considerable leverage. Table 2.6 shows that commercial banks in different size classes have a ratio of equity capital to total assets that averaged a little more than 10 percent and banks in no size group had a capital ratio as high as 15 percent in 2004. Such high leverage ratios are seen as facilitating the role played by commercial banks in the payments system.

The role of commercial banks in the payments system derives from their twin roles as distributor of currency (paper money and coin), and as producer and servicer of demand deposits. Currency and demand deposits are the community's principal *means of payment* and *media of exchange*, and are the major components of the money supply. Commercial banks link the central bank with the many millions of money users.

This nexus reflects on our second point, that commercial banks are central because of their role in monetary policy. The central bank seeks to stabilize economic activity by controlling the money available to support that activity. Hence, if inflation threatens, for example, the Federal Reserve will restrain the growth of money and drive up interest rates. Restricting growth of the money supply reduces the availability of bank credit to commercial banks, thereby lowering the volume of the loans they make and driving up the loan interest rates. This is the way commercial banks transmit monetary policy and fulfill their stabilizing role. We will revisit this issue in Chapter 3.

TABLE 2.6 FDIC–Insured Commercial Banks in 2004

Asset Size	Return on Assets (in Percent)	Return on Equity (in Percent)	Equity Capital (in Percent of Assets)
Less than $100 million	0.99	8.46	11.52
$100 million to $1 billion	1.28	12.88	10.00
$1 billion to $10 billion	1.46	13.48	10.90
More than $10 billion	1.30	14.24	9.95
Total	1.31	13.82	10.10

Source: FDIC Quarterly Banking Profile, December 2004.

2. Virtually every country in the world has a central bank charged with managing the money supply, acting as lender of last resort, protecting the integrity of the financial system, and other related chores. The Federal Reserve is the central bank of the United States. Counterparts in other countries include the Bank of England, the Bank of Japan, the Bundesbank in Germany, to mention a few. The European Central Bank acts as the central bank for the European Union. Central banks are typically government owned, but the Federal Reserve has a peculiar hybrid structure reflecting a populist ambivalence toward concentrations of economic and financial power, particularly in the hands of the government. Thus, the Federal Reserve is nominally independent of government and privately owned, but as a practical matter it is neither.

In playing this role in the conduct of monetary policy, a commercial bank acts both as a broker and as a QAT, providing all of the services listed in Figure 2.2 except management expertise. A typical commercial bank's balance sheet and its sources of revenues and expenses are shown in Tables 2.7 and 2.8, respectively.

U.S. commercial banks are regulated by the Federal Reserve System, the Office of the Comptroller of the Century (OCC), and the Federal Deposit Insurance Corporation (FDIC) at the federal level and by state banking authorities at the state level. The major regulations that banks are subject to are discussed in Appendix 2.2.

Commercial banks have many things in common with other depository institutions, but are distinguished by their above-mentioned role in the payments system, by the diversity of their assets, and by their ownership structure. Other depositories, such as savings institutions (often called thrifts) and credit unions, have traditionally had more narrowly specialized asset portfolios—residential mortgages and consumer credit comprise the bulk of their assets, respectively, although these distinctions

TABLE 2.7 Hypothetical Balance Sheet for a U.S. Bank

Assets			
Cash and Due			$125
Securities Held			170
Federal Funds Sold			50
Loans:			
Real Estate		160	
Commercial and Industrial		220	
Consumer		110	
All Other		120	
Less Unearned Income:		−12	
Allowances for Possible Loan Losses			
Total Loans			598
Other Assets			57
Total Assets			$1,000
Liabilities and Equity			
Liabilities:			
Deposits			
Domestic	$661		
Foreign	119		
Total Deposits		$780	
Federal Funds Purchased		80	
Other Liabilities		80	
Total Liabilities			940
Subordinated Notes and Debentures			5
Equity Capital:			
Preferred and Common Stock		10	
Surplus		20	
Undivided Profits and Reserves		25	
Total Equity Capital			55
Total Liabilities and Equity			$1,000

TABLE 2.8 Major Revenue and Expense Items for a Bank

Revenues	Expenses
–Interest on loans and marketable securities	–Interest on deposits
–Fees on loan commitments and other contingent claims	–Wages
–Fees on cash management and other transactions services	–Operating expenses, including occupancy
	–Deposit insurance premia
	–Taxes
	–Provisions for loan losses

have been blurring and are almost irrelevant now for the most part. Commercial banks, as their name suggests, were also specialized lenders at one time, but they have evolved to the point that the largest of them hold a great variety of earning assets, including working capital, trade and term financing for businesses, residential and commercial mortgages, consumer loans, automobile loans, loans to sovereigns (governments), structured financing for corporate buyouts, and still more exotic credit instruments. In addition, commercial banks perform risk-shifting functions through their sale of standby letters of credit, swaps, and other financial guarantees. These are called "contingent claims" and have many of the attributes of ordinary insurance contracts.

The ownership structure of commercial banks is also notably different from other depository institutions. Commercial banks alone are *all* shareholder owned. Thrifts are substantially mutual, and credit unions are exclusively mutual, that is, they are owned by their depositors (a discussion of mutual organizations appears in the next chapter). In this era of galloping globalization, it is noteworthy that American commercial banking still reflects peculiarly American concerns. Interestingly, many of these idiosyncrasies are shared by the Japanese despite profound cultural differences. This is because Japanese banking was patterned after U.S. institutions following World War II. Indeed, our financial system probably shares more in common with Japan's than with those of our other major trading partners in Europe and the Americas.

For example, numerous major banks in Europe—including the largest in France—and in Mexico have traditionally been government rather than privately owned.[3] Commercial banks in the United States and Japan also historically tended to be more narrowly restricted in their activities (this distinction, like so many others, has eroded substantially under the pressures of global competition) and the consequent deregulation. For examples, Germany's "universal" or "haus" banks are permitted to engage in all manner of insurance and investment banking, as well as the many activities traditionally permitted American commercial banks. Such activities were traditionally proscribed for American commercial banks, but these restrictions have since been removed with the dismantling of the Glass-Steagall Act.

In addition to being more narrowly restricted functionally, commercial banks in the United States have also been geographically confined. Until recently, commercial banks in the United States could not operate in more than one state, with minor

3. Mexico has been "privatizing" its banking system, as have the former Communist–bloc countries like Poland and Romania. China's banks are still government owned and controlled, but have private sector minority owners too.

TABLE 2.9 International Comparison of Bank Concentration and Profitability

	Top Five Banks' Assets (In Percent of All Banks' Assets)		*Number of Branches (In Thousands)*		*Pretax Profits (In Percent of Average Assets)*[b]	*Net Interest Margin (In Percent of Average Assets)*[b]
Country	**1990**	**2003**[a]	**1990**	**2003**[a]	**2003**	**2003**
Canada	83	87	8.7	10.4	1.00	1.99
France	52	45	25.7	26.2	0.59	0.80
Germany	17	22	45.3	38.2	−0.12	0.81
Italy	24	27	17.7	29.9	1.03	2.82
Japan	42	42	24.7	22.7	−0.47	1.21
Netherlands	74	84	8.0	3.7	0.65	1.62
Spain	38	55	35.2	39.4	1.29	2.45
Sweden	70	90	3.3	2.0	0.77	1.44
Switzerland	54	80	4.2	2.7	0.59	0.97
U.K.	49	41	19.0	12.9	1.22	1.96
U.S.	13	24	72.8	84.8	2.10	3.21

[a] For France, Germany, Italy, Japan, Sweden and the U.K. 2002.
[b] Profitability calculations relate to major banks only. Number of banks included: Canada (5), France (7) Germany (9), Italy (6), Japan (11), Netherlands (3), Spain (5), Sweden (4), Switzerland (5), U.K. (9), U.S. (12).
Source: Bank for International Settlements 75th Annual Report Thrifts.

exceptions. Indeed, in many states, commercial banks could not operate more than one office. This may seem quaint, but these Americanisms gave rise to over 30,000 independently chartered commercial banks at their peak, about 90 years ago. Markets were Balkanized and entry was restricted, reflecting America's populist fear of economic power concentrations, especially when such power resided in the major eastern urban centers where the country's largest financial institutions were headquartered. Also reflected in these policies (laws) was America's fear of large-scale bank failures. Recall that these practices predate federally sponsored deposit insurance, which originated in the 1930s. Populist sentiments trace back to frontier America when the West sought cheap money, manufacturers, and transport. The Eastern establishment wanted sound money and sound banks, along with market-determined prices for railroad services and manufactured goods.[4] While these sentiments now seem outdated, and deregulation in 1994 now permits interstate branching, the United States still had over 7,600 banks in 2004. This means the U.S. has far more banks than other countries. Even with the recent trend toward consolidation, the U.S. retains a relatively fragmented banking market with many independent, albeit a few large ones. Table 2.9 provides an international comparison.

Thrifts

Savings and loan associations (S&Ls) and mutual savings banks (MSBs), collectively referred to as thrifts, or savings institutions, are depository institutions that were specially chartered to extend residential mortgage finance. Traditionally their assets

4. The railroads, in particular, enjoyed market power, owing to the paucity of substitute conveyances, and this served as one basis for protracted regional conflicts.

are primarily home mortgages, although their asset mix has been changing to include other assets in recent years. Thrifts have traditionally had even lower capital ratios than banks, as shown in Table 2.10, although regulation following the thrift failures in the 1980s has resulted in thrift capital ratios moving up, and actually being higher in 2004 than the average bank capital ratio (in Table 2.6).

S&Ls were chartered for the purpose of specializing in consumer savings accounts and residential mortgage loans. They came into existence to encourage thrift and to allow people to purchase homes at a time when banks were loathe to finance home mortgages. They started out as small informal mutuals, and despite the fact that many have converted into stockholder-owned institutions, mutual S&Ls abound. They were regulated by the independent Federal Home Loan Bank Board (FHLBB) and insured by the Federal Savings and Loan Insurance Corporation (FSLIC). In 1989, the FHLBB was dissolved and replaced by the Office of Thrift Supervision (OTS) under the control of the Treasury, and the OTS is now the main regulatory body for S&Ls. The FDIC now provides federal deposit insurance for S&Ls. The financial intermediation services provided by S&Ls are similar to those provided by commercial banks, but with different emphasis.

Mutual savings banks, as their name indicates, are cooperatively owned. Like S&Ls, they too invest mostly in mortgage loans and marketable securities. They are a few hundred in number and most are located in the Northeast and the Northwest of the United States. MSBs managed to distance themselves, at least for a time, from the misfortune of the savings and loans in recent years;[5] they were regulated by the FDIC rather than the now-defunct FHLBB and the FSLIC. The MSBs held less risky assets and operated with less financial leverage than the S&Ls, and had the good fortune to be located away from some of the worst real estate markets of the 1980s—Texas, Oklahoma, Louisiana, and Colorado. MSBs were nevertheless damaged by the inflation-induced loss of core deposits, the consequent emergence of interest-rate risk, and the asset-quality problems of the later '80s. Earlier proud pillars of the industry like The Bowery Savings Bank of New York and The Philadelphia Savings Fund Society were forced into humiliating restructurings, emerging as shareholder-owned shadows of their former selves.

The 1989 FIRREA (Financial Institutions and Regulatory Reform Act) legislation, which did away with the FSLIC and the Federal Home Loan Bank Board, and folded the savings and loan federal deposit insurance fund into the FDIC, further

TABLE 2.10 Key Statistics Regarding Federally Insured Savings Institutions

	1990	1995	2000	2004
Number of Institutions	2,815	2,030	1,589	1,345
Net Worth to Total Assets	4.11%	7.84%	8.68%	11.18%
Return on Assets	−0.35%	0.70%	0.91%	1.17%
Return on Equity	−7.65%	9.00%	11.63%	12.79%
Net Income	−$3.8 billion	$5.4 billion	$8.0 billion	$14.0 billion
Net Worth	$67.5 billion	$86.1 billion	$103.6 billion	$189.1 billion
Total Assets	$1,260 billion	$1,026 billion	$1,223 billion	$1,692 billion

Source: Office of Thrift Supervision 2004 Fact Book.

5. For engaging accounts of these, see Martin Mayer (1990), and James R. Adams (1990).

weakened the distinctions between MSBs and S&Ls.[6] As a practical matter, the distinction between MSBs and their cousins, the S&Ls, has been lost in a deluge of asset-quality problems. They are now less undifferentiated parties to the thrift industry implosion, estimated to cost taxpayers upwards of $250 billion in present value terms as of mid-1990, although subsequent estimates put the cost at around $100 billion[7].

The thrift fiasco was a large financial disaster. A whole industry with over thousands of firms and trillions of dollars was devastated. The industry seems to have recovered, however. Although their numbers have diminished, thrifts continue to operate successfully. Whether any will remain dedicated housing lenders for long is questionable, however. Table 2.10 provides further information on thrifts. It indicates that the financial condition of the industry is improving through time, although the numbers of thrifts has been declining through time.

In diagnosing the thrift industry collapse, some point to flaws in the deposit insurance system, particularly the failure to relate deposit insurance premiums or capital requirements to the risk assumed by the thrifts. The deposit insurance contract provided inappropriate risk-taking incentives to thrift managers. Nevertheless, for four decades, it worked like a charm. These issues will be taken up in later chapters.

Credit Unions

Like thrifts, credit unions specialize in consumer savings and are mutuals. Those forming a credit union must share *a common bond*, that is, they should be employed by the same organization. The credit union must be involved in borrowing and lending to its members. The homogeneity in borrower base facilitates the credit union's control of credit risks, but limits potential diversification. As of year-end 2004, there were 5,572 credit unions in the United States.

A credit union's liabilities consist mainly of consumer deposits, and its assets are comprised mainly of consumer loans; real estate mortgages to members; loans to other credit unions, MSBs, and S&Ls; and government and corporate securities. Federally chartered credit unions are regulated by the National Credit Union Administration (NCUA), which also provides deposit insurance. State-chartered institutions can purchase NCUA deposit insurance as well. The services provided by a credit union include transactions services, screening, origination, monitoring, funding, guaranteeing, and liquidity creation. Like their other depository brethren, credit unions have low capital-to-total assets (stated as "reserves to assets") ratios, as shown in Table 2.11.

6. Although S&L deposit insurance is administered by the FDIC along with commercial bank deposit insurance, separate insurance funds are maintained. Members of FDIC, including MSBs, are insured by the Bank Insurance Fund (BIF), whereas former FSLIC members are insured by the Savings Association Insurance Fund (SAIF). Are these beltway acronyms mnemonic or ironic?

7. Loss estimates ranging from $1/4 trillion upwards were obtained by assuming long-term financing and adding in the interest cost. Described as a "bailout," the loss was merely a spectacular example of a governmental guarantee program run amok. We have many such government programs in housing, health, education, agriculture, and similar, if less spectacular fiascos have visited these programs. The Farm Credit Administration failure of the 1980s is an illustration.

TABLE 2.11 Federal Credit Unions-Significant Ratios 1990–2004

	1990	1995	2000	2004
Number of Institutions	8,539	7,329	6,336	5,572
Reserves to Assets	4.0%	4.3%	4.5%	3.7%
Reserves and Undivided Earnings to Assets	75	10.2	11.6	11.0
Reserves to Loans	6.2	6.9	6.6	6.0
Loans to Shares	70.4	70.8	78.0	72.6
Operating Expenses to Gross Income	35.7	42.3	44.8	59.7
Salaries and Benefits to Gross Income	15.0	19.2	20.2	27.0
Dividends to Gross Income	55.7	42.6	41.7	23.1
Yield on Average Assets	10.6	8.1	8.3	5.8

Source: 1995 and 2004 Annual Reports, National Credit Union Administration, Washington, DC.

Nondepository Intermediaries

The primary focus of this book is deposit-taking financial intermediaries, and most specially commercial banks. However, commercial banks are members of a vast and diverse financial services industry with overlapping markets and regulatory jurisdictions. These jurisdictional and competitive relationships condition behaviors with regard to pricing, output, attitudes toward risk, and just about every other facet of the business of financial intermediation. Therefore, we shall spend the next few pages sketching some of the more interesting members of this fascinating industry.

Venture Capitalists

Most fledgling entrepreneurs, are unable to obtain bank financing. They go instead to venture capitalists. Many prominent firms, including Apple and Federal Express, began with funding from venture capitalists. Venture capitalists typically provide *both* capital and expertise that allow entrepreneurs to convert ideas into commercial ventures.

Venture capital funding is normally in the form of structured financing, including both equity and convertible debt, rather than just the loans that banks provide. The salient features of a venture capital contract are as follows:[8]

1. The entrepreneur cannot ‘‘walk away’’ after obtaining financing and negotiate with another financier (no *de novo* financing).
2. The entrepreneur may be relieved of control of the firm by the venture capitalist unless the firm's performance meets some minimum requirement (‘‘performance requirement’’).
3. If the entrepreneur is relieved of control, he is paid a fixed amount independent of his demonstrated skill and subsequent cash flows of the firm; that is, he is bought out by the venture capitalist (‘‘buyout’’ option for the venture capitalist).
4. If control remains with the entrepreneur, *both* the venture capitalist and the entrepreneur receive equity payoffs (‘‘earnout’’ arrangement).

8. The discussion in this section is based on Chan, Siegel, and Thakor (1990). There is a large literature on venture capital. See, for example, Hellmann and Puri (2000).

Why do venture capital contracts have these features? To understand this, consider the parties involved in these transactions. First, the entrepreneur often is either an engineer (who knows the manufacturing technology of the product to sell) or a marketing expert, but he is often inexperienced in managing *all* facets of a business. Venture capitalists, while they are not necessarily intimately familiar with the production or marketing technique of the products their clients want to produce, usually have considerable management expertise and a nose for "troubleshooting" based on experience in financing and managing numerous ventures. Thus, venture capitalists possess two attributes that entrepreneurs need: financial capital and management expertise.

Typically, when the partnership between a venture capitalist and an entrepreneur commences, neither is completely sure of the entrepreneur's management ability. The initial period is one of learning for both parties. If the firm's performance were to indicate that the entrepreneur lacked sufficient management ability, it would be efficient to replace the entrepreneur with the venture capitalist to avail of the latter's managerial skills. Although such a passage of control may be the best thing for the firm, it is not obvious that the entrepreneur would be eager or even willing to relinquish control of the firm. That is, the entrepreneur's attachment to the venture he thought of and started may stand in the way of implementing the best *ex post* plan for the firm. To prevent this, both parties could agree *ex ante* to an explicit clause in the venture capital contract that allows for an orderly transfer of control. This would also benefit the entrepreneur *ex ante*, as the venture capitalist's recognition of the possibility that he can buy out the entrepreneur and take control of the firm if things go really badly will improve the terms of the initial financing received by the entrepreneur. In this regard, it is also important for the venture capitalist to have an equity claim in the firm. This not only gives the venture capitalist a more active voice in management even when the entrepreneur is in control, but also provides the venture capitalist with all the ownership incentives to invest managerial expertise in the firm and to counsel the entrepreneur.

We can now see why banks may be unwilling to lend to most entrepreneurs. Because banks don't possess managerial expertise in running young, nonfinancial firms, and also because regulation prevents them from doing so except during short, transitional periods following borrower default, the performance requirements and buyout options used by the venture capitalist are not available to the bank. Thus, if the entrepreneur turns out to be a poor manager and the business fails, the bank can do little to revive and nurture it back to success. It would simply be left holding the assets of the firm (assuming the entrepreneur defaults on his loan) as collateral of possibly dubious value. Hence, the same entrepreneur is generally more risky to the bank than to the venture capitalist.

Why don't banks hire management consultants to assist entrepreneurs in their fledgling businesses? They could. However, a management consultant would be merely an agent of the bank and thus the bank would confront *moral hazard* in motivating the consultant. A venture capitalist avoids this moral hazard by *combining* management expertise and financing into one entity. Alternatively, banks could hire the talent needed to advise entrepreneurs, as do the venture capitalists. However, because of the short-term nature of the bank's liabilities and their government guarantee, equity-type claims have traditionally been viewed as inappropriate for banks.

Banks are at less of a disadvantage relative to venture capitalists in dealing with well-established borrowers. With this more stable clientele, the bank's superior ability to access the capital markets and its ability to avail itself of deposit insurance and the discount window give it a distinct advantage. Not surprisingly, then, for firms where the manager's ability to shepherd his organization through the more vulnerable early phase is not a critical issue, bank loans tend to be preferred to venture capital. Moreover, in these cases, managers will prefer to avoid the possibility of having to relinquish control to the venture capitalist at some future time. Thus, the intermediation services provided by venture capitalists include screening and certification, funding, monitoring, management expertise, and liquidity creation and transformation.

Finance Companies

Most of the important finance companies originated as narrowly focused trade finance subsidiaries of large nonfinancial companies. Examples include General Electric Capital, General Motors Acceptance Corporation (GMAC), and FOMOCO (Ford Motor Company's finance subsidiary). Others have independent origins, including factoring specialists.[9] Finance companies lend to consumers for auto and home purchases, as well as other purposes, and to businesses for a wide range of applications. These intermediaries usually specialize in processing riskier credits, and most of their lending is done on a secured basis, that is, with collateral, unlike commercial banks that lend on both unsecured and secured bases. There are three basic types of finance companies: sales finance companies that make car and appliance loans; personal finance companies, which make small personal loans (for example, for debt consolidation); and business finance companies, which make commercial loans and leases. The intermediation services provided by these finance companies include screening, origination, funding, and claims transformation.

Finance companies typically fund themselves by selling commercial paper. Indeed, the most compelling difference between commercial banks and finance companies is in their primary sources of funding. Because the commercial banks are substantially funded by governmentally insured deposits, they are invested with a special public interest and are subject to pervasive regulation. The finance companies do not have access to subsidized funds and are not subject to regulatory restrictions, proscriptions, examinations, and supervision.

The commercial paper sold by finance companies is an unsecured general obligation of the issuer and has a fixed maturity of less than 9 months. Most often, the maturity of commercial paper is shorter than 6 months at date of issue. Commercial paper is typically sold in large denominations and is rated by specialized credit-rating agencies. Because the paper is unsecured, issuers are usually compelled to purchase a dedicated (back-up) loan commitment in order to obtain a favorable credit rating. Back-up loan commitments are sold by commercial banks expressly for the purpose of providing the commercial paper issuer with the funds to redeem its paper in case rolling over the maturing paper proves to be infeasible. The back-up commitment from the bank ensures the commercial paper issuer's ability to redeem its paper, conditional only on the bank's performance on its loan commitment.

9. Factors provide working capital and/or collection services by purchasing and/or servicing the accounts receivable of nonfinancial firms. Factoring is an early form of asset-backed lending, done with or without recourse.

The commercial paper market is notoriously fragile. Macroeconomic shocks have been known to paralyze commercial paper issuance. In addition, the fortunes of a particular borrower may preclude use of the commercial paper market. Thus, the back-up loan commitments, or ‘‘back-up lines,’’ are critically important to the lender, especially in light of the unsecured status of commercial paper.

Finance companies are an illustration of the evolution of competition between regulated and largely unregulated segments of the financial services industry. The captive finance subsidiaries, the most important players in this market segment, grew out of trade credit provided by larger, better-rated nonfinancial corporations. Having developed the expertise necessary to underwrite the credit of their customers and having established secure sources of funding, why wouldn't one offer these financial services to the community beyond one's customer base? This is the logic that drove GMAC from the exclusive financing of its own auto sales to becoming the largest home-mortgage servicer in the United States. It is the same logic that drove Ford Motor Credit into the savings and loan business and General Electric Capital into virtually every facet of banking (with the notable exception of deposit taking), including the ownership of Kidder Peabody, a failed major investment bank in the United States.

Trade credit, the driver in this story of horizontal and vertical integration, arises moreover as the natural accompaniment to trade between parties with widely disparate access to the credit markets.[10] Consider General Electric (GE), a large and well-rated company that sells industrial equipment to smaller, less well-known companies. In periods of credit stringency, GE's customers are crowded (rationed) out of the credit market well before GE, and this reduces the demand for GE's products. In order to smooth the cyclicality of demand for its goods, GE will borrow in order to provide its customer with uninterrupted access to credit. This will stabilize and also increase the demand for GE's nonfinancial output. Enhanced revenues and decreased cost should ensue, the latter owing to more predictable production runs and smaller inventories.

Trade credit is a natural complement to trade in nonfinancial goods and services whenever traders have different degrees of access to capital markets. It illustrates a very basic attribute of banking, namely the negligible natural barriers to entry. Thus, absent regulatory restrictions, one would expect to see a steady flow of new financial intermediaries entering and others departing (failing or merging) as the industry adjusts to changes in the demand for its services. Hence, those that specialize in the provision of financial services can expect competition from their own clients who enjoy the advantage of being largely unregulated, but must therefore borrow in the open market without the benefit of government subsidies.

The market share of finance companies, measured in terms of asset size, is a small fraction of that of commercial banks,[11] but this probably understates the importance of finance company competition, especially for the money-center and super regional banks that typically serve the same customers, middle-market companies and the larger consumer markets.

10. Although less common, there is no reason why trade credit cannot flow from buyer to seller. Wal-Mart, Costco and Home Depot are much larger and often more creditworthy than their suppliers. One would expect these retailing giants to offer credit to reduce the likelihood of supply interruptions and to benefit from the reduced production cost their suppliers would experience as a result of regularized production and reduced inventories.

11. See the Federal Reserve's Flow of Funds Accounts.

Insurance Companies

Private sector life and health insurance companies manage trillions of dollars of assets. Property and casualty insurers control hundreds of billions of dollars in assets. Together, the insurers are slightly less than half the size of the commercial banking industry. As in the case of thrifts, many insurance companies are organized as mutuals (cooperatives) rather than as shareholder-owned institutions. Some key statistics pertaining to the life insurance industry in the U.S. are provided in Table 2.12. As is evident, life insurance firms invest the premiums they collect in a wide variety of assets, including real estate.

Insurance companies hold many of the same kinds of assets found on the balance sheets of commercial banks, but insurer assets are financed for the most part with *contingent* liabilities. That is, the insurance company's liabilities become current (or terminate, in the case of annuities) upon the occurrence of some prespecified event, the timing or realization of which is inherently uncertain when the insurance contract is written. Insurance is written against a large variety of contingencies. Life insurance companies typically contract against the expiration of life or the realization of health care needs.[12] Property and casualty insurers write policies against: i) damage or loss of physical or intellectual property, including the loss of income or extraordinary expenses associated with the property damage or loss, ii) liability, iii) health care needs, and iv) surety. Surety contracts guarantee third-party contractual performance. Examples include fidelity, construction and bail bonds, and also standby letters of credit that are a mainstay of the commercial banking business. Standby letters typically guarantee the repayment of third-party debt.

For example, A might be vaguely interested in extending credit to B, but may not be entirely sure about repayment prospects. A then may request that B arrange a standby letter of credit with her bank, insurance company or other credible financial guarantor. In exchange for the payment of an appropriate insurance premium,[13] the guarantor will accept the risk of repaying A's loan to B in the event B fails to do so. This kind of financial guarantee is commonly written by commercial banks, property and casualty (multiline) insurers, and even "pure" financial guarantors (monoline insurers) who do nothing more than guarantee performance of third parties under debt contracts.

The most striking differences between banks and insurance companies are found on the liability sides of their respective balance sheets. Wherever the liabilities of banks change, often instantaneously and at the sole discretion of depositors, insurance liabilities change on the occurrence of events largely uncontrollable by the claimant. In addition, the duration of life insurance liabilities, in particular, is much longer than that of commercial bank deposit liabilities. Thus, life insurers

12. Life and health insurance are genteel euphemisms that support marketing efforts. It is more difficult, to be sure, to sell death and illness insurance.

13. Robert Mehr explains the origin of the term "premium" in insurance that directly links the insurance business to commercial banking: "If a Greek shipowner planned a voyage to bring cargo from a foreign land, he would borrow the necessary money by pledging his ship as collateral. The contract provided that if the ship failed to return to port intact, the lender would have no claim against the shipowner. This type of contract [called bottomry] became common throughout maritime countries... The interest charged on these contracts included a sum in addition to that normally charged for the loan to compensate the lender for writing insurance [accepting credit risk] to cover the safety of the voyage. This additional amount, logically, was called a premium, and to this day the consideration paid for insurance is still referred to as a premium." *Fundamentals of Insurance*, 2nd edition, Irwin, p. 13.

TABLE 2.12 U.S. Life Insurance Companies-Significant Ratios 1990–2002

Item	1990	1995	2000	2002
U.S. Companies[1]	2,195	2,079	1,269	1,171
Income	402.2	528.1	811.5	734.0
Life Insurance Premiums	76.7	102.8	130.6	134.5
Annuity Considerations[2]	129.1	152.4	306.7	269.3
Health Insurance Premiums	58.3	90.0	105.6	108.7
Investment and Other	138.2	176.9	268.5	221.5
Payments under Life Insurance and Annuity Contracts	88.4	227.6	375.2	301.3
Payments under Life Insurance Beneficiaries	24.6	34.5	44.1	48.2
Surrender Values under Life Insurance[3]	18.0	19.5	27.2	32.9
Surrender Values under Annuity Contracts[3,4]	n.a	105.4	214.0	142.9
Policyholder Dividends	12.0	17.8	20.0	21.0
Annuity Payments	32.6	48.5	68.7	55.0
Matured Endowments	0.7	1.0	0.6	0.6
Other Payments	0.6	0.9	0.6	0.6
Health Insurance Benefit Payments	40.0	64.7	78.8	78.7
Assets	1,408	2,144	3,182	3,380
Government Bonds	211	409	364	481
Corporate Securities	711	1,241	2,238	2,266
(Percent of Total Assets)	50	58	70	67
Bonds	583	869	1,241	1,475
Stocks	128	372	997	791
Mortgages	270	212	237	251
Real Estate	43	52	36	33
Policy Loans	63	96	102	105
Other	110	133	204	244
Interest Earned on Assets (In Percent)[5]	8.89	7.41	7.05	5.38
Obligations and Surplus Funds[6]	1,408	2,144	1,241	1,475
Policy Reserves	1,197	1,812	2,712	2,507
Annuities[7]	798	1,213	1,841	1,550
Group	516	619	960	570
Individual	282	594	881	980
Supplementary Contracts[8]	17	25	34	14
Life Insurance	349	511	742	833
Health Insurance	33	63	96	111
Liabilities for Deposit-Type Contracts[9]	18	20	21	364
Capital and Surplus	91	151	188	202

n.a. = Not Available
1. Includes life insurance companies that sell accident and health insurance in 2000 and 2002.
2. Excludes certain deposit-type funds from income due to codification in 2002.
3. "Surrender values" include annuity withdrawals of funds in 2000 and 2002.
4. Excludes payments under deposit-type contracts in 2002.
5. Net rate.
6. Includes other obligations not shown separately.
7. Excludes reserves for guaranteed interest contracts in 2002.
8. Includes(excludes) reserves for contracts with and without life contingencies in 1994 and 2000 (2002).
9. Policyholder dividend accumulations for all years.
Source: U.S. Census Bureau, Statistical Abstract of the United States: 2004–2005.

can and do hold longer-term assets than commercial banks. This difference in duration may be the most fundamental difference between banks and life insurers (the liabilities of property and casualty insurers tend to be shorter than those of life insurers). Life insurers and pension funds are allegedly the two largest private-sector pools of long-term money.[14]

In order to distinguish insurance from the kind of risk-shifting that takes place through the purchase (sale) of a financial futures contract or an option or a "swap" contract, insurance is commonly defined to involve some application of the law of large numbers. Thus, insurance requires some pooling of risks among independent events to avail of diversification and make it easier to price such risks. It is difficult for investors to avail of such diversification themselves because of the bulky unit size of some claims. Again, a strong analogy between insurance and banking emerges: diversification enables banks to manage credit and withdrawal (interest rate) risks, and individuals' limited wealth and access to credit markets limits the potential for "homemade" diversification. The intermediation services provided by insurance companies include screening and certification, origination, funding, monitoring, guaranteeing, and claims transformation.

Pensions

Along with life insurance, private pension funds accumulate the long-term liabilities that are capable of funding the durable assets so critical to real capital accumulation. In earlier years, when bank deposits were subject to interest rate ceilings and competition was more restrained, banks and thrifts too were capable of making 7- to 10-year fixed-interest-rate business loans and even 30-year fixed-rate mortgages with acceptable levels of interest rate risk. The shortened duration of deposits, however, has rendered banks and thrifts less able to provide long-term credit. To be sure, banks and thrifts offer longer-lived loans, but the interest rates on them are typically variable. These sequences of short-term loans provide the borrower with no certainty regarding future for longer-term credits, and this has elevated the importance of the pension funds and life insurance.[15]

Private pension funds, along with mutual funds, are the only two major financial intermediaries to have steadily growing market share since 1953. Forty years ago, private pension funds had 5 percent, or approximately one-tenth of the commercial banks' market share, but by 1990 the banks had fallen to 30 percent and the pension controlled two-thirds of the banks' share. By virtue of their size, momentum, and the extended duration of their liabilities, pension funds have become a major domestic private-sector influence on capital formation.

14. The careful reader will note that this distinction is easily overdrawn in that policy loans can be made against nonterm life insurance policies at the owners discretion. Moreover, life insurance can "lapse" as a result of the insured's decision not to make timely insurance payments. A second nuance relates to the distinction between discretionary withdrawals of depositors and the presumably uncontrollable random events that trigger insurance claims. Most states of nature that trigger insurance claims are subject to some human influence. This ability to affect the insured contingencies is referred to as moral hazard (see Chapter 1).

15. Likewise, the departure of banks from term lending has elevated the importance of financial futures, options, and swaps, which are risk-shifting financial contracts that permit the borrower to dispose of part of the unwanted interest rate risk of an indexed loan.

The liabilities of defined-contribution pension funds are actuated upon retirement or death of their members, at which time the member's claim is paid out as a lump sum or used to purchase an annuity. In the case of defined-benefit plans, the retirement fund pays a prescribed annuity to the claimant upon retirement. In the time interval between contributions and the termination of the funds' liability to the claimants, the contributions are invested in a wide variety of assets, everything from real estate and other equities to Treasury debt. These investments are constrained by federal legislation (ERISA), which defines the responsibilities of pension fiduciaries. The key intermediation services provided by pension funds are guaranteeing and claims transformation.

Pension funds are being called upon increasingly to play a role in corporate governance as representatives of their millions of beneficiaries. Historically, the pensions have been passive investors, but issues like the composition of boards of directors, executive compensation, potential conflicts of interest of executives involved in buyouts and many other issues vitally affect current and future retirees with investments in corporate America. The problem is that the pension fund managers typically hold investments in many hundreds of corporations—indeed, many adopt consciously passive strategies of cloning the stock indexes (purchasing securities that behave like the averages)—and they are simply not staffed adequately to participate in the affairs of individual corporations. This, however, is increasingly unacceptable to pension participants and the community at large as more instances of corporate abuse are widely chronicled. It seems inevitable that the guardians of America's pension assets will be forced to become more active in corporate affairs, and this will no doubt affect corporate governance in the future. A factor that potentially affects these dynamics is that, like deposits, pensions are now federally insured.[16]

Mutual Funds

Along with pension funds, mutual funds have been major market-share winners over the past 40 years. Essentially a post-World War II phenomenon, mutual funds have risen from an inconsequential share of the intermediation market in 1950 to achieve a 6 percent market share in 1990, and a 17 percent market share in 2003 (measured based on total assets). Its significant growth can also be gleaned from the penetration of mutual funds among U.S. households, which increased from 22 percent in 1990 to 43 percent in 2004 (see Table 2.2). By 2005, these variegated investment vehicles had grown to about 60 percent of the size of commercial banks and larger than pension funds, insurance companies, and savings institutions.

Mutual funds come in two basic varieties: open- and closed-end. Closed-end funds have a pre-established number of shares and the fund's initial resources typically are not augmented with the subsequent sale of shares. A closed-end fund is typically traded as a single security on organized exchanges, for example, the New York Stock Exchange, and its shares are priced directly in the market like the shares of any other company. As a consequence, the market price of closed-end fund shares can deviate, often widely, from the liquidation value of the securities they hold. Open-end funds operate on very different rules. Their shares are continuously liquidated and augmented by a specialized management company that offers shares for cash, and cash

16. An interesting part of this dynamic is that defined contribution plans are displacing defined benefit plans.

for shares at net asset value (NAV). NAV is the estimated liquidation or market value of the fund's assets divided by the number of shares the fund has outstanding. Thus, unlike closed-end fund shares, the prices of open-end fund shares cannot deviate from the value of underlying assets.

The open-end funds have given rise to large specialized fund management companies, like Fidelity, Scudder, Vanguard, and Dreyfus. Each of these manages and markets a wide range of different funds, each of which is defined in terms of specific investment objectives. For example, were you to phone the appropriate 800 number, Vanguard would be pleased to inundate you with literature describing the numerous different funds it manages. These investment companies earn their keep by levying fees against the funds it manages. The funds, of course, are owned by their investors. Were you to consult the financial pages of any major newspaper, you would find a section headed mutual funds wherein you could find the NAV of any of the numerous mutual funds managed by Merrill Lynch, or any of the very large number managed by Fidelity. These larger mutual fund companies typically have tens of billions of dollars under management. The key financial intermediation services provided by mutual funds include transactions services, screening, and certification.

There's nothing terribly new about mutual funds, except their explosive growth in recent decades. There are at least three reasons for the current popularity of the funds. First, money-market mutual funds, which were introduced in the 1960s, rapidly became the instrument of choice for circumventing Regulation Q deposit interest rate ceilings. As inflation accelerated in the 1970s and market interest rates soared, the spreads between these rates and deposit rates gaped ever wider. The bloated opportunity cost of holding bank deposits increased the appeal of money-market funds. The rest is history! Despite the competitive disadvantage of operating without a government guarantee, the mutual funds grew spectacularly, underscoring that there are limits to what the public is willing to pay for governmental deposit insurance.

By and large, the money-market funds were managed conservatively, and some even restricted themselves to holding direct debt of the U.S. government. More commonly, the funds held negotiable large-denomination certificates of deposit of world-class banks, commercial paper, bankers' acceptances, mortgage and other asset-backed securities, and government agency debt. Almost all of these assets were less than one year to maturity, and the funds traded at a constant one dollar per share.

Moreover, the money-market funds are sustained by implicit guarantees of their managers. In at least three cases, management companies made good on asset losses in order to protect their own reputations and the viability of the money funds they managed. For example, Value Line manages a money-market fund that held the commercial paper of Integrated Resources, a company that defaulted on its debt. Rather than reflect this loss in its money-market fund, which almost certainly would have meant the fund's demise, Value Line management bought the Integrated Resources commercial paper from its money-market fund at par. Notably, there was no legal or even moral obligation to protect the fund's investors, but the action was presumably motivated by the desire to maintain and build upon Value Line's reputation in managing money-market funds. Clearly, the money-market funds offered a compelling package of substitutes for the governmental deposit guarantee. Low-risk investment strategies, combined with implicit guarantees of reputable management companies, and substantially higher yields permitted the money-market funds to ravage the bank and thrift deposit markets and enjoy meteoric growth.

The second and third reasons for the recent growth of mutual funds are less dramatic, but nevertheless noteworthy. In recent decades, the public has gradually become persuaded of the improbability of consistently "beating" the stock market. A sea of research, much of it academic, has demonstrated that over most extended spans of time asset managers do less well than the widely watched stock market indices, for example, Dow Jones, and Standard and Poor's. The reasons are numerous and complex, but the facts seem plain. The widespread acceptance of this idea has had a profound effect on investment behavior, and in particular it has led to the idea that if you can't beat the averages, you can do no better than to buy the averages. Buying the averages is known as passive investment. This is done by purchasing a portfolio of securities that behave like (clone) the averages. Since this strategy typically requires holding a substantial number of securities, it is often infeasible for smaller wealthholders, and uneconomic for most. However, mutual funds can provide such a service at low cost. Thus, the popularity of passive investment strategies provides a second reason for the recent growth of mutual funds.

Finally, the past six decades have witnessed the much-heralded globalization of financial markets. Many investors believe it is as important to diversify across economies (currencies) as it is to diversify across industries. Furthermore, diversification across economies has been massively simplified in recent decades, as regulatory and tax barriers have been dismantled. However, information about foreign investment opportunities is still relatively expensive. Hence, the mutual fund has become the instrument of choice for investing abroad. Many "country funds" are closed-end and listed on the New York Stock Exchange, but there are also many open-end funds that specialize in countries and regions of the world. To mix a metaphor, as the pie of foreign indirect investment has grown larger, the bologna of specialization among funds has been sliced ever thinner.

Hedge Funds

In contrast to most mutual funds, hedge funds are actively managed funds that pursue nontraditional investment strategies. A hedge fund is a private investment pool subject to the terms of an investment agreement between the sponsor of the fund and its investors. They take both long and short positions in a variety of instruments – equities, fixed income securities, currencies, etc. – to achieve the highest return commensurate with the fund's objectives. Although the hedge fund industry has traditionally been far less regulated than mutual funds, that gap was closed in 2004, when hedge funds were required to register under the Investment Advisers Act. This act allows the SEC to inspect all hedge fund advisers for approval purposes. Moreover, hedge funds are now subject to many of the same requirements as mutual fund advisers.

Differences between hedge funds and mutual funds persist, however. While mutual fund sales charges and fees are subject to regulatory limits, there are no limits on the fees hedge fund advisers can charge.

Also, mutual funds are restricted in their ability to leverage against the value of securities in their portfolio, whereas leveraging and other higher-risk investment strategies are commonplace for hedge funds. In fact, hedge funds originally came into existence to invest in equity securities and use leverage and short selling to hedge the exposure of the portfolio to stock price movements. Finally, while any investor can open a mutual fund account with $1,000 or less, a minimum investment of $1 million or more is typically required to become a hedge fund investor.

Investment Banking

Investment banks, like Merrill Lynch, Salomon Brothers, and Morgan Stanley, specialize in the design and issuance of financial contracts. They often perform the brokerage function of bringing buyers and sellers of securities together. The key intermediation services they provide are transactions services, financial advice, screening and certification, origination, issuance, and guaranteeing.

Investment banks and bankers have allegedly played a central role in the corporate corruption of the 1980s, the late 1990s, and the early part of the 21st century. Who can forget the scandalous accounts of commercial and investment banking activities narrated in bestsellers like *Bonfire of the Vanities, Barbarians at the Gate, Wall Street, The Predator's Ball, Liar's Poker, The Big Fix, The Greatest Ever Bank Robbery, Conspiracy of Fools*, *The Smartest Guys in the Room*, and *Confessions of a Wall Street Analyst*? Who can forget the convulsive implosion of Drexel Burnham Lambert and the spectacle of Michael Milken confessing to seven felony counts? And in the late 1990s, we had other corporate scandals like World Com and Global Crossing that also brought investment banks into the public limelight.

These citadels of entrepreneurial hubris bore their traditional substitutionary and complementary relationship to the commercial banks. The investment banks' marketing of equities complemented the commercial banks' provision of loans. At the same time, however, the investment banks sold fixed-income securities, including bonds and commercial paper, that competed directly with commercial bank loans. Similarly, the investment banks aggressively marketed money-market funds in competition with commercial banks, while at the same time they brokered deposits to the banks.

This multifaceted and ambivalent relationship, sometimes symbiotic and sometimes subversive, was a major theme of the 1980s that expressed itself darkly in the thrift debacle and subsequent disarray in commercial banking. In thinking about these tragedies, note that the very existence of Wall Street, as we know it, is the result of questionable legislation of the 1930s (Glass-Steagall Act) that erected a high but not altogether impermeable wall of separation between commercial banks and investment banks. This legislation created investment banking in its singular American incarnation. No other major country, with the possible exception of Japan, had the kind of separation found in the United States. The European model is that of "universal banks" that bridge the chasm between the two forms and permit rationalization of structures dictated by the economics of the business.

Glass-Steagall created two banking systems. Commercial bankers had subsidized deposits, but restricted asset choice. The investment banks were without deposit subsidies, but had great freedom on the asset side of the balance sheet, and protection from commercial bank competition in equity markets. This permitted the investment bankers to selectively attack the commercial banks' niches of profitability, forcing them into ever riskier endeavors in order to justify the capital committed to commercial banking. The intricate and exquisitely contradictory relationship between commercial banks and investment banks is a product of American history. Their functions are substantially overlapping. The forms and instruments employed often differ only at a superficial level. Moreover, the investment banks are as much a creation of 1930s banking legislation as the commercial banks. With the dismantling of the Glass-Steagall Act and the passage of the Gramm-Leach-Bliley Act, this artificial separation between commercial and investment banking has been finally eliminated.

The Role of the Government

To this point, we have sketched the major players in the world of financial intermediation. Probably the most important intermediaries to add to this list are the vast government enterprises that routinely provide a wide variety of financial services. These would include the Old Age, Survivors and Disability Insurance, Workers' Compensation, Medicare, the housing agencies (Federal National Mortgage Corporations of FNMA or "Fannie Mae," Federal Home Loan Mortgage Corporation or FHLMC or "Freddie Mac," and the Government National Mortgage Association or GNMA or "Ginnie Mae") Farm Credit Administration, Small Business Administration, Student Loan Marketing Association (or "Sally Mae"), and flood insurance programs of the Agriculture Department. And the list goes on!

Annual payments to the federal government's Old Age, Survivors, Disability Insurance, and Medicare programs are twice the *assets* of the largest commercial bank in the United States, and about one-sixth the assets of the entire commercial banking industry. Without doubt, the U.S. government is far and away the largest financial services provider in the country and arguably in the world.

Financial Intermediaries on the Periphery

Gambling

Prominent on the periphery of the financial intermediation universe is the glamorous world of legal and illegal gambling. Some deny that gambling is a financial service, but this seems a quibble. The bookmaker is as much a broker as the trader of options and financial futures. The naysayers argue that gambling *creates* risk, whereas insurance dissipates and redistributes pre-existing risk. But whether the gambling relates to a manufactured uncertainty (for example, a horse race or roulette) or to some pre-existing natural process (for example, the number of live pups your neighbor's dog will whelp), seems incidental. The production of uncertainty is logically separable and incidental to the gambling.

The more meaningful distinction between insurance and gambling is that the former involves the exchange of a certain cost (the premium) for *relief* from an uncertain liability, whereas the latter is the exchange of a certain cost (say the price of a lottery ticket) for an uncertain future receipt. The bookmaker would just as soon wager on tomorrow's mean temperature as on the three-digit numbers generated by tomorrow's horse races. It matters not whether the bet is hedging or speculating, nor does it matter what process generates the uncertainty.[17] The bookmaker merely fills a market niche, one usually scorned or illegal. The difference between the bookmaker and the insurance agent may well be that one is legal and the other is not, but at a deeper level the insurer sells alleviation from risk to those ill-equipped to bear it, whereas the bookmaker sells risk to those who find it welfare-improving. In this latter sense, both are brokers, and possibly qualitative asset transformers too. The bookmaker is a financial intermediary in the same sense as the insurance agent or underwriter.

Pawnbrokers

Also on the periphery we have "bankers" to the poor and the excluded (who perforce are high-risk borrowers). The major participants in these market niches are pawnbrokers and loan sharks, the former legal and the latter not usually. As of 1991, there were in the United States approximately three times as many pawnbrokers (about 6,900) as savings and loan associations.[18] Pawn loans are typically small, say \$50–\$100. Most of these loans are for a few weeks, sometimes months, and all are secured with merchandise (jewelry, electronics, musical instruments, guns, and the like) with a resale value roughly twice the debt. All-in interest rates range from high to astronomical, and can be as high as 25–30 percent per month in states without interest rate ceilings. In 2004, it is estimated that there were 15,000 pawnbrokers in the U. S.[19]

Pawnbroking is a traditional form of asset-backed (secured) lending. The lender typically prefers to be repaid rather than taking ownership and liquidating the collateral (this is because the failure to repay usually ruptures a valuable customer relationship), but the creditworthiness of the borrower is rarely at issue (the pawnbroker rarely has the information necessary to form an intelligent judgment, except perhaps in cases of longtime customers). The loan is made entirely on the basis of the borrower's collateral. Default rates between 10 and 30 percent are common. The intermediation services provided by pawnbrokers include origination, funding, and market completeness.

The pawnbroker industry began to stagnate in the late 1990s with the rise of payday and title lending alternatives, which are discussed below.[20]

Payday Lending

Payday lenders did not operate as a formal industry until the early 1990s. Prior to this time, most payday lenders were check cashers who made payday loans as a casual extension of their core business. By 2004, there were 12,000 payday lenders in the United States,[21] with major pawn chains having also entered the business.

17. If one views the bookmaker as inherently dishonest, one might prefer to gamble on a process subject to human influence, perhaps his own (moral hazard). But such an assumption about bookmakers seems gratuitous and beside the point.

The gambling enterprise is so vast that we find it done in both the public sector (lotteries) and in the private sector. In the latter, there are legal expressions (parimutuel betting, both on- and off-track and casinos) and illegal expressions (bookmaking and the "numbers game").

18. See Caskey (1991).

19. See Fass and Francis (2004).

20. See Caskey (2003).

21. See Barr (2004). The discussion below is based on Barr (2004).

Payday lenders provide unsecured short-term loans to customers. The loan arises in one of two ways. One is the traditional payday loan transaction, in which the borrower writes a postdated (or undated) personal check to a lender, the lender makes a loan equal in amount to the check minus finance charge. The lender holds the check before either depositing it, or receiving cash repayment directly from the borrower, usually on the borrower's payday. The second is a variant of the traditional transaction, in which no check is written, but the borrower signs an authorization that permits the lender to debit his bank account on a future date for the amount of the loan plus the finance charge. The typical loan term is two weeks.

The payday lending industry has grown to approximately 12,000 firms in 31 states and DC. In 2000, payday lenders made about 65 million loans to 8–10 million households, totaling $8–$14 billion in loan value, and generating over $2 billion in revenue. The industry reports gross margins of 30–45 percent of revenue, with losses at 1–1.3 percent of receivables, and return on investment of 24 percent.

Title Lenders

Title lenders are similar to payday lenders, the difference being that title lenders make secured loans rather than unsecured loans. That is, instead of holding a check or debit authorization until payday, title lenders hold collateral against the loan. Typically, $250 to $1,000, and the value of the associated collateral is typically three times as much.

The title lending industry is essentially an extension of the pawnbroker industry. The two differences between them are as follows: First, a pawnbroker keeps physical possession of the collateral until the loan is repaid, whereas a title lender may permit the collateral to physically rest with the borrower during the loan term and repossess it only upon default. Second, title loans are typically larger than pawn loans. These two differences, however, are not economically important for distinguishing between these two types of lenders in terms of the brokerage and QAT functions served by them. That is, payday lenders and title lenders serve essentially the same economic functions as pawnbrokers.

Like loans extended by pawnbrokers, payday loans and loans made by title lenders tend to have very high interest rates, often exceeding 25 percent *per month*, for an annual percentage rate (APR) of 300 percent. The title loan industry originated in the southeastern United States and has spread to other states like Missouri, Illinois and Oregon. In some states, an upper limit of 30 percent annual interest rate was imposed, which essentially eliminated the industry there.

Loan Sharks

Whereas the pawnbroker lives on the edge of respectability (see the splendid movie of the same title, with Rod Steiger), loan sharks live beyond the pale. Dates on loansharking are understandably sketchy, but these financial intermediaries play a prominent role in providing credit in support of both legal and illegal enterprises.[22] The President's Crime Commission in 1967 asserted that loan-sharking was the second most important activity of organized crime.

22. For a fascinating description of the business, see Reuter and Rubinstein (1982), and Haller and Alvitti (1977).

A definitional note will help to clarify much confusion. If by loan-sharking we mean *all* illegal lending, loan-sharking will include an amorphous hodgepodge of lenders who violate usury laws. More useful, it would seem, is to think of loan sharks as lenders who can credibly make illegal or socially unacceptable threats of violence and intimidation in connection with collections. The availability of this singular and extralegal collection technology explains why this financial service is provided by criminal elements, why interest rates on such loans tend to be high, and why their clientele are typically desperate borrowers with few alternatives. The legality of their activities aside, loan sharks serve economic functions that are similar to those of payday and title lenders. In fact, some refer to payday lenders as "legal loan sharks."

Reuter and Rubinstein describe three kinds of loans made by loan sharks. Short-term small loans of under $1,000 were made on a weekly six-for-five basis. Loans of $1,000 or more, called "knockdowns," would call for 12 weekly repayments of $100. A third type of loan, usually for larger amounts, called a "vig" loan, would call for weekly interest payments of one-half to 3 percent with the principal returned *in toto* at termination of the loan.

The same authors also describe the fairly common use of collateral, but this would seem to be an anomaly, unless the credit is to be used for illegal purposes. A properly secured loan would obviate the need for, or usefulness of extralegal intimidation. Hence, the borrower should be able to borrow from any asset-based lender such as a finance company or a pawnbroker at considerably lower interest rates than those quoted by loan sharks. However, legitimate lenders could be expected to avoid lending to felons, or for projects known to be illegal.

Apparently, a substantial fraction of the loans made by loan sharks are to bookmakers down on their luck. It would not be surprising to learn that much credit also goes to finance illicit drug and stolen goods inventories. But the less glamorous side of loan-sharking must be lending to the fringes of society without the collateral to offer a pawnbroker or finance company. To these unfortunates, the loan shark offers a service that no law-abiding institution, short of a charity, can provide. Whatever the moral considerations, loan sharks are nevertheless an indispensable part of the financial services industry. They are bankers to the poor, the forgotten, and to those living outside the law.

Conclusion

This chapter has provided a selective survey of the major and more interesting members of the financial services industry. We used our description of commercial banks and thrifts to also sketch the financial environment. The deposit revolution continues to reshape deposit-dependent institutions, and very likely this portion of the financial services industry will be fundamentally restructured in the next five years. Either deposit insurance and regulation will be reconfigured or depository F.I.s will continue to lose market share to the less regulated segments of the industry.

Major competitors for commercial banks and thrifts include insurance companies, finance companies, pensions, and mutual funds. The linkages among these segments, the cutting edge of competition, are described in the respective sections on each. The theme is one of commonality and similarity; differences among segments of the industry are seen as legal, artificial, and exaggerated. And of course, one can never forget the government (". . . where does the gorilla sleep?") as a member of this gigantic industry.

Finally, we addressed a collection of important and often neglected financial intermediaries on the periphery of the industry. Included in this collection is the woolly world of gambling—public and private, legal and illegal—and the shadowy backwaters of pawnbroking, payday and title lending, and loan-sharking. All have their assigned roles, based on the law and technology, in processing risk and information and in allocating credit. Each serves as broker and/or asset transformer, and absent the more bizarre actions associated with the criminal aspects of some of these activities, each makes the market work more effectively, thereby increasing the economic pie available to be shared among all.

Review Questions

1. Given below is an excerpt from "A Friendly Conversation." Who do you agree with? Provide a thorough discussion of the theoretical and empirical underpinnings of your opinion.

   Appleton: Absolutely! I believe that when you cut through all the bull, the essential role of banks is to act as "lot breakers" and provide simple transactions service. I can't write checks against a T-bill, so I need a bank.
   Butterworth: Alex, I couldn't disagree more. Everything that I've read suggests that banks *are* special. Your proposal would destroy a key ingredient of the process by which society allocates capital from savers to investors.
   Moderator: It looks to me like we have a *fundamental* disagreement: Why do we have banks and what do they really do?
   Appleton: What's to disagree? Ask anybody and they'll say that banks are there to borrow and lend money.
   Moderator: That's obvious, but it hardly settles the issue, does it, Alex? After all, borrowing and lending are *not* services in themselves, but rather the *visible outcomes* of banks' production of financial services. The question is: What are these less transparent financial services that banks and other financial intermediaries produce? You say that the services are purely transactional, while Beth claims they are much more.

2. Discuss what is meant by brokerage and asset transformation. What factors determine the value of brokerage services?
3. List five distinct types of financial intermediaries, explain what they do, and provide a comparison/contrast of the basic intermediation services they provide.
4. Find information on capital-to-total-assets ratios for several nonfinancial firms and compare them to those for financial firms. Why the differences?
5. From the information in Table 2.6, what can you conclude about the risk in holding a representative bank's equity compared to that in holding equity in a diversified market portfolio?

Appendix 2.1 Measurement Distortions and the Balance Sheet

The balance sheet perspective on financial intermediation provided in the Introduction is suggestive but stylized and therefore incomplete.

The balance sheet, for banks as well as other entities, is an accounting statement that states the values of the firm's cash flows as of some specified date. In principle, the listing of assets is exhaustive and if the valuations are done properly, the remainder or net assets constitutes a sensible (unbiased) estimate of the firm's capital or net worth. However, in practice, assets are occasionally omitted (arbitrarily valued at zero), while others are improperly valued. Indeed, the principles of valuation vary across categories of assets, so that the net worth is often difficult, if not impossible, to interpret.

For example, if reputational capital is purchased, it is carried on the balance sheet at its depreciated purchase price. Called "goodwill," this asset is usually written off according to some arbitrary schedule chosen by auditors and/or other interested parties such as governmental regulatory agents. If, on the other hand, the firm chooses to develop a reputation, as opposed to purchasing an existing one, generally accepted accounting principles will accord the reputational capital zero value. Accountants defend this inconsistent treatment with reference to their "conservatism." However, from an economist's viewpoint, the practice distorts or biases balance sheets. Moreover, in a world of costly capital and information, the incentive to develop reputation is weakened by the asymmetric accounting treatment.

Now consider earning assets such as loans and securities. The accounting convention is that assets held for "trading" purposes must be marked to market, whereas those assets held for "investment" may be carried at adjusted historical cost. If the latter assets perform unexceptionally, the assets often are carried at original cost.[23] Moreover, there is no unambiguous basis for distinguishing between trading and investment motives, so the auditors exercise their discretion. This notion of valuing assets at cost seems bizarre to those naïve enough to think of the balance sheet as a description of the firm's financial condition, but many of the investment assets are *not* traded in active markets and it is therefore difficult to value them at arbitrary points in time, like December 31 and June 30. This is a systemic rather than an aberrant problem, in the sense that the *raison d'etre* of banks is to serve as repositories for those assets without active secondary markets. This is how the bank produces liquidity! But accurate point estimates of the values of such assets are inherently difficult to come by and auditors are understandably loathe to oblige, given the litigious inclinations of their disparate clienteles.

The issue of Generally Accepted Accounting Principles (GAAP) versus current (or market) value accounting has been in the forefront of the ongoing debate. However, it is difficult to know what current value accounting would mean in markets with wide bid-ask spreads. Forced to do current value accounting, the auditors might insist on interval rather than point estimates, or perhaps refuse to certify the accuracy of their estimates. Would the market then be better informed? Would managers display less pathological behavior? Perhaps! Noisy, unbiased estimates may well be superior to less noisy, but biased alternatives.

The valuation problem, it should be noted, expresses itself on both sides of the balance sheet. Core deposits, for example, are treated as investment rather than trading assets, and they are carried at par, cost, or redemption value. Thus a dollar

23. Loans have occasionally been written down by examiners despite unexceptional performance. This typically happens when the loan has an interest reserve account that temporarily services the credit, but the financial condition of the borrower has deteriorated to the point where its ability to service the loans after the interest reserve has been exhausted is brought into question. Hence the oxymoronic "performing nonperformers."

of deposits is invariably a dollar of liability. Note, however, that when banks are sold, their deposits typically command a premium. The buyer is willing to pay (typically between 1 and 6 percent) for the deposits. Why? Because deposits are inexpensive as a source of funding. They embody subsidy or "rent" deriving from underpriced deposit insurance and restricted entry into banking. But then, shouldn't the valuation of deposits reflect these rents or subsidies? Doesn't the failure to account for them overstate the bank's liabilities and understate its net worth? This dubious accounting practice may overstate the stability of the F.I.s net worth. This distortion gave rise to much of the "hidden" capital in banking, thought to be so important in reducing banks' appetites for risk taking.

In any case, the bank balance sheet reflects a complex mix of disparate valuation practices that confound the best efforts at interpretation. Some argue that current value accounting would do the community a disservice by adding volatility to reported financial results. The counterargument is that GAAP data knowingly mislead and compromise the integrity of the system that produces such data.

Appendix 2.2 Guide to Federal Reserve Regulations

Regulation A – Loans to Depository Institutions

Regulation A governs borrowing by depository institutions at the Federal Reserve discount window, which is available to any depository institution that maintains transaction accounts or nonpersonal time deposits. The purpose of the discount window is to provide short-term liquidity, typically overnight, for depository institutions in need. Government securities are usually used as collateral.

Regulation B – Equal Credit Opportunity

Regulation B prohibits creditors from discriminating improperly against credit applicants, establishes guidelines for gathering and evaluating credit information, and requires written notification when credit is denied. The regulation prohibits creditors from discrimination against applicants on the basis of age, race, color, religion, national origin, gender, marital status, or receipt of income from public assistance programs. As a general rule, creditors may not ask (on applications) the race, color, religion, national origin, or gender of applicants. Exceptions apply in the case of residential mortgage applications. In addition, if the application is for individual, unsecured credit, the creditor may not ask the applicant's marital status. Creditors also may not discriminate against applicants who exercise their rights under the federal consumer credit laws.

The regulation also requires creditors to give applicants a written notification of rejection of an application, a statement of the applicant's rights under the Equal Credit Opportunity Act, and a statement either of the reasons for the rejection or of the applicant's right to request the reasons. Creditors who furnish credit information when reporting information on married borrowers must report information on the names of each spouse.

The regulation establishes a special residential mortgage credit monitoring system for regulatory agencies by requiring that lenders ask for and note the race, national origin, sex, marital status, and age of residential mortgage applicants.

The regulation covers all credit transactions (unlike other regulations that may cover only consumer credit), with some modifications applicable to certain classes of transactions.

Regulation C – Home Mortgage Disclosure

Regulation *C* requires certain mortgage lenders to disclose data regarding their lending patterns. The regulation carries out the Home Mortgage Disclosure Act of 1975, providing citizens and public officials with data to help determine whether lenders are meeting the credit needs of their communities and complying with fair lending laws.

The regulation applies to banks, savings and loans, credit unions, and certain mortgage companies that have offices in Metropolitan Statistical Areas and assets of greater than $10 million or originated at least one hundred home purchase loans. These institutions must publicly disclose data on mortgage loans that they originate or purchase and also on applications for such loans. In many instances, the race or national origin, gender, and income of the applicant must be reported as well as the locations of the property and the type of loan.

Regulation D – Reserve Requirements

A reserve requirement is a stipulation that the bank keep a minimum fraction of its deposits and Eurocurrency liabilities as liquid assets, either vault cash or deposits held at the Federal Reserve. Regulation D imposes uniform reserve requirements on all depository institutions with transaction accounts or nonpersonal time deposits, defines such deposits and requires reports of deposits and Eurocurrency liabilities to the Federal Reserve.

Regulation E – Electronic Fund Transfers

Regulation *E* establishes the rights, liabilities, and responsibilities of parties in electronic fund transfers (EFT) and protects consumers using EFT systems. Regulation *E* specifies rules for the solicitation and issuance of EFT cards, governs consumer liability for unauthorized EFTs (for example, from lost or stolen cards), requires institutions to disclose certain terms and conditions of EFT services, provides for documentation of electronic transfers (on periodic statements, for example), sets up a resolution procedure for errors on EFTs and covers notice of crediting and stoppage of preauthorized payments from a customer's account.

Regulation F – Limitations on Interbank Liabilities

This regulation prescribes standards to limit the risks that the failure of a depository institution would pose for an insured institution. In particular, it limits a bank's interday credit exposure to an individual correspondent to no more than 25 percent of the bank's total capital, unless the correspondent is at least adequately capitalized.

Regulation G – Disclosure and Reporting of CRA-Related Agreements

Regulation *G* implements the Gramm-Leach-Bliley Community Reinvestment Act (CRA) Sunshine Requirements provisions. It generally requires nongovernmental entities or persons and insured depository institutions or affiliates that are parties to certain written agreements made in fulfillment of the CRA to make the agreements available to the public and to the relevant supervisory agency, and file annual reports concerning those agreements with the relevant supervisory agency. In addition to describing factors related to the fulfillment of the CRA, Regulation *G* also provides criteria for determining when an agreement is a "covered agreement," thus triggering the disclosure and annual reporting requirements of the regulation.

Regulation H – Membership Requirements for State-Chartered Banks

Regulation *H* defines the requirements for membership of state-chartered banks in the Federal Reserve System; sets limitations on certain investments and requirements for certain types of loans; describes rules pertaining to securities-related activities; establishes the minimum ratios of capital to assets that banks must maintain and procedures for prompt corrective action when banks are not adequately capitalized; prescribes real estate lending and appraisal standards; sets out requirements concerning bank security procedures, suspicious-activity reports, and compliance with the Bank Secrecy Act; and establishes rules governing banks' ownership or control of financial subsidiaries.

Regulation I – Member Stock in Federal Reserve Banks

Regulation *I* requires each bank joining the Federal Reserve System to subscribe to the stock of its District Reserve Bank in an amount equal to 6 percent of the member bank's capital and surplus. Half the total must be paid on approval. The remainder is subject to call by the board of governors. A 6 percent dividend is distributed on paid-in portions of Reserve Bank stock. Ownership of stock does not carry with it the usual attributes of control and financial interest. The stock is not transferable and cannot be used as collateral.

Regulation J – Check Collection and Funds Transfer

Regulation *J* establishes procedures, duties, and responsibilities among Federal Reserve Banks and (a) the senders and payers of checks and other items, and (b) the senders and recipients of wire transfers of funds. Regulation J provides a legal framework for depository institutions to collect checks and settle balances through the Federal Reserve System. The regulation specifies terms and conditions under which Reserve Banks will receive items for collection from depository institutions and will present items to depository institutions. Along with Regulation CC, Regulation J establishes rules under which depository institutions may return unpaid checks through Reserve Banks, as well as terms and conditions under which Reserve Banks will receive and deliver transfers of funds over Fedwire, the Federal Reserve's wire transfer system, from and to depository institutions.

Regulation K – International Banking Operations

Regulation *K* governs the international banking operations of U.S. banking organizations and foreign banks in the United States. It also governs the operations of Edge-Act corporations, the international operations of U.S. banks and bank holding companies, the interstate banking and certain nonbanking activities of foreign banks in the United States, the operations of bank-affiliated export trading companies, and certain international lending practices of bank holding companies and state member banks.

Regulation L – Interlocking Bank Relationships

Regulation *L* avoids restraints on competition among depository organizations by restricting the interlocking relationships that a management official may have with a depository organization. The regulation prohibits a management official of a depository institution or depository institution holding company from serving simultaneously as a management official of another depository organization if the two organizations are unaffiliated, very large, or located in the same local area.

Regulation M – Consumer Leasing

Regulation *M* implements the consumer leasing provisions of the Truth in Lending Act. It applies to leases of personal property for more than 4 months and for a total contractual obligation not exceeding $25,000 for personal, family, or household use. It requires leasing companies to disclose in writing the cost of a lease, including security deposit, monthly payments, license, registration, taxes, and maintenance fees and, in the case of an open-end lease, whether a balloon payment may be applied. It also requires written disclosure of the terms of a lease, including insurance, guarantees, responsibility for servicing the property, standards for wear and tear, and any option to buy.

Regulation N – Relationships With Foreign Banks

Regulation *N* is internal to the Federal Reserve System. It governs relationships and transactions among Reserve Banks and foreign banks, bankers, and governments, and describes the role of the Board of Governors in these relationships and transactions. The regulation governs the relations of Reserve Banks with foreign banks and foreign governments and provides for special supervision of these activities by the board. The regulation provides that the Reserve Banks must receive the prior permission of the board before negotiating with foreign banks or foreign governments. In addition, Reserve Banks may not enter into any agreements, contracts, or understandings with any foreign banks of foreign governments without prior permission of the board.

Regulation O – Loans to Executive Officers of Member Banks

Regulation *O* restricts credit extended by a member bank to its executive officers, directors, and principal shareholders and their related interests. Further, the regulation imposes reporting requirements relating to credit extended by a correspondent bank to a member bank's executive officers and principal shareholders and their related interests.

Regulation P – Privacy of Consumer Financial Information

Regulation *P* governs the treatment of nonpublic personal financial information about consumers who obtain financial products or services primarily for personal, family or household purposes from any financial institutions for which the Federal Reserve Board had primary supervisory authority (including state member banks and bank holding companies). It specifies that financial institutions must provide a clear and conspicuous notice that accurately reflects privacy policies and practices to its customers and specifies the information to be included.

Regulation Q – Interest on Deposits

Regulation Q prohibits member banks from paying interest on demand deposits and prescribes rules for advertising deposits. Many interest rate restrictions have by now been phased out, under the Depository Institution Deregulation and Monetary Control Act of 1980.

Regulation R – Interlocking Relationships Between Securities Dealers and Member Banks (Rescinded in 1996)

Regulation S – Reimbursement for Providing Financial Records

Regulation *S* establishes the rates and conditions for reimbursement to financial institutions for providing records to a government authority.

Regulation T – Margin Credit Extended by Brokers and Dealers

Regulation *T* governs credit extensions by securities brokers and dealers, including all members of national securities exchanges. The regulation applies to broker-dealers and all national securities exchange members. In general, they may not extend credit to their customers unless the loan is secured by margin securities.

The term *margin securities* includes: any equity security listed on or having unlisted trading privileges on a national securities exchange; any security listed on NASDAQ; any nonequity security; any foreign margin stock; any debt security convertible into a margin security; and mutual funds.

Regulation U – Margin Credit Extended by Banks and Persons Other Than Brokers and Dealers

Regulation *U* governs extension of credit by persons other than brokers and dealers for purchasing and carrying margin securities. The regulation applies to entities other than brokers and dealers that are extending credit that is secured, directly or indirectly, by margin stock. Any time a loan is made in an amount that exceeds $100,000 in which a margin stock serves as collateral, the lender must have the customer execute a purpose statement regardless of the use of the loan. The margin requirements imposed by the regulation apply if the loan is both margin-stock secured and is for the purpose of purchasing or carrying margin stock. Certain exceptions exist for specified special purpose loans to broker-dealers, for loans to qualified employee stock option plans, or for loans to plan lenders.

Regulation V – Fair Credit Reporting

Regulation *V* specifies that there be proper disposal of consumer information obtained by member banks of the Federal Reserve System (other than national banks) and their respective operating subs, branches and agencies of foreign banks, and commercial lending companies owned or controlled by foreign banks.

Regulation W – Transactions Between Member Banks and Their Affiliates

Regulation *W* establishes certain quantitative limits and other prudential requirements for loans, purchases of assets, and certain other transactions between a member bank and its affiliates.

Regulation X – Borrowers Who Obtain Margin Credit

Regulation *X* extends the provisions of Regulations *T* and *U* (governing extensions of credit for purchasing or carrying securities in the United States) to certain borrowers and to certain types of credit extensions not specifically covered by those regulations.

Regulation Y – Bank Holding Companies and Change in Bank Control

Regulation *Y* governs the bank and nonbank expansion of bank holding companies and to the divestiture of impermissible nonbank interests. Regulation *Y* also governs the acquisition of a bank by individuals. Under the Bank Holding Company Act of 1956, as amended, a bank holding company is a company that directly or indirectly owns or controls a bank. The regulation contains presumptions and procedures the board uses to determine whether a company controls a bank. The regulation also explains the procedures for obtaining board approval to become a bank holding company and procedures to be followed by bank holding companies acquiring voting shares of banks of nonbank companies. It also governs the establishment and activities of financial holding companies.

Regulation Z – Truth in Lending

Regulation Z prescribes uniform methods of computing the cost of credit, disclosure of credit terms, and procedures for resolving errors on consumer credit accounts. Consumer credit is generally defined as credit offered or extended to individuals for personal, family, or household purposes, where the credit is repayable in more than four installments or for which a finance charge is imposed. The major provisions of the regulation require lenders to:

- provide borrowers with meaningful, written information on essential credit terms, including the cost of credit expressed as an annual percentage rate (APR);
- respond to consumer complaints of billing errors on certain credit accounts within a specified period;
- identify credit transactions on periodic statements of open and credit accounts;
- provide certain rights regarding credit cards;
- provide good-faith estimates of disclosure information before consummation of certain residential mortgage transactions;
- provide early disclosure of credit terms to consumers interested in adjustable rate mortgages (ARMs) and home equity lines of credit; and
- comply with special requirements when advertising credit.

Regulation AA – Consumer Complaint Procedures

Regulation *AA* establishes consumer complaint procedures and defines unfair or deceptive acts or practices of banks in connection with extensions of credit to consumers. Under the regulation, a consumer complaint concerning either an alleged unfair or deceptive act or practice, or an alleged violation of law or regulation by a state member bank will be referred to the appropriate federal agencies.

Regulation BB – Community Reinvestment

Regulation *BB* implements the Community Reinvestment Act (CRA) and is designed to encourage banks to help meet the credit needs of their communities. Under Regulation BB, each bank office must make available a statement for public inspection indicating, on a map, the communities served by that office and the type of credit the bank is prepared to extend within the communities served. The regulation requires each bank to maintain a file of public comments relating to its CRA statement. The Federal Reserve Board must assess the bank's record in meeting the credit needs of the entire community, including low- and moderate-income neighborhoods, and must take account of the record in considering certain bank applications. In addition, recent amendments to the CRA will require public disclosure of a bank's CRA rating and the CRA performance evaluations.

Regulation CC – Availability of Funds and Collection of Checks

Regulation *CC* implements the Expedited Funds Availability Act (EFA) and governs the availability of funds and the collection and return of checks. Regulation *CC* establishes availability schedules, as provided in the EFA, under which depository institutions must make funds deposited into transaction accounts available for withdrawal. The regulation also provides that depository institutions must disclose their funds availability policies to their customers. In addition, Regulation *CC* establishes rules designed to speed the collection and return of checks and imposes a responsibility on banks to return unpaid checks expeditiously. The provisions of Regulation *CC* govern all checks, not just those collected through the Federal Reserve System.

Regulation DD – Truth in Savings

Regulation *DD* requires depository institutions to disclose the terms of deposit accounts to consumers. The regulation applies to consumer deposit accounts offered by depository institutions (except credit unions, which are governed by rules of the National Credit Union Administration). Regulation *DD* enables consumers to make informed decisions about accounts at depository institutions by requiring those institutions to: provide consumer account holders with written information about important terms of an account, including the annual percentage yield; provide fee and other information on any periodic statement sent to consumers; use certain methods to determine the balance on which interest is calculated; comply with special requirements when advertising deposit accounts.

Regulation EE – Netting Eligibility for Financial Institutions

Regulation EE aims to enhance efficiency and reduce systemic risk in financial markets. It defines financial institutions to be covered by statutory provisions that validate netting contracts, thereby permitting one institution to pay or receive the net, rather than the gross, amount due, even if the other institution is insolvent.

Regulation FF – Obtaining and Using Medical Information in Connection With Credit

Regulation FF establishes final rules creating exceptions to the statutory prohibition against obtaining or using medical information in connection with determining eligibility for credit (effective April 1, 2006).

Other Important Regulations

Banks are also subject to *capital requirements, portfolio restrictions*, and *branching restrictions*. Each bank is required to keep at least a certain fraction of its assets as capital. Recent changes in capital requirements, involving the international harmonization of capital standards, have led to capital requirements being linked to the default risks of the bank's assets. Consequently, there is a different percentage requirement against each category of assets on the bank's balance sheet. Moreover, capital is also required to be held against certain categories of off-balance sheet items (that is, those claims that do not appear on the bank's balance sheet), such as standby letters of credit and some loan commitments. If capital requirements are not satisfied, banks are subject to restrictions on their activities until capital is adequately refurbished. The box below lists the corrective actions required under the FDIC Improvement Act of 1991 if a bank's capital falls below target levels.

Banks were also subject to strict restrictions on the compositions of their asset portfolios. A U.S. bank generally could not hold equity in a corporation and could not undertake investment banking and insurance activities, with minor exceptions. These restrictions were dismantled with the passage of the Gramm-Leach-Bliley Act in 1999. See the box below.

FDIC Improvement Act of 1991: Key Corrective Actions Required

(A) Well-Capitalized Banks (Capital no less than 10 percent of total assets)
- Capital distributions that could result in undercapitalization are prohibited.
- Management fees that could lead to undercapitalization are prohibited.
- If the bank is found to be engaging in unsafe practices, it may be reclassified as adequately capitalized or undercapitalized.

(B) Adequately Capitalized (Capital less than 8 percent–10 percent)
- Capital distributions are management fees that could result in undercapitalization.
- Limited regulatory monitoring.

(C) Undercapitalized (Capital less than 8 percent)
- Capital distributions are restricted.
- Management fees are prohibited.
- Regulatory monitoring is required.
- Asset growth is restricted.
- New branch openings and acquisitions are restricted.
- Additional discretionary regulatory actions are possible.
- A capital restoration plan must be filed.

(D) Significantly Undercapitalized (Capital less than 6 percent capital)
- Sales of additional stock may be required.
- Institution could be merged with a better-capitalized institution.
- Transactions with affiliates may be restricted.
- Asset growth and interest rates on deposits may be restricted.
- "Risky" activities may be curtailed.
- Election for new directors may be ordered.
- Directors and officers who held office immediately prior to the undercapitalization may be dismissed.
- Restrictions may also be placed on various aspects of executive compensation.

(E) Critically Undercapitalized (Capital less than 2 percent)
- Without prior FDIC approval, the following activities are prohibited: HLT credit, charter/bylaw amendments, material accounting changes, excessive compensation/bonuses, and higher interest on new or renewing liabilities.
- No principal or interest payments on subordinated debt (outstanding after July 15, 1991) are allowed.
- A conservator or receiver will be appointed for the institution by regulators.

The Financial Services Modernization Act of 1999 aka The Gramm-Leach Bliley Act of 1999

- The Gramm-Leach-Bliley Act of 1999 facilitates affiliation between banks and securities firms by repealing Sections 20 and 32 of the Glass-Steagall Act.
- The act authorizes bank holding companies and foreign banks that meet eligibility criteria to become financial holding companies, thus allowing them to engage in a broad array of financially related activities.
- The act also provides for the functional regulation of financial holding companies, protects nonpublic customer information held by financial institutions, alters supervision related to the Community Reinvestment Act (CRA), and makes various other regulatory changes.

References

Adams, James Ring, *The Big Fix: Inside the S&L Scandal: How an Unholy Alliance of Politics and Money Destroyed America's Banking System*, 1990, John Wiley & Sons, Inc.

Barr, Michael S., "Banking the Poor," *Yale Law Journal* 21, 2004.

Banker's Diary & Guide 1993, Warren Gorham and Lamont, Division of Research Institute of America, Inc.

Boyd, John, and Mark Gertler, "U.S. Commercial Banking: Trends, Cycles, and Policy," paper prepared for the 1993 *NBER Macroeconomics Annual*, February 1993.

Carkey, John, "Fringe Banking a Decade Later," working paper, 2003.

Caskey, John P., "Pawnbroking in America: The Economics of a Forgotten Credit Market," *Journal of Money, Credit and Banking* 23 (9), February 1991, 85–99.

Chan, Yuk-Shee, Daniel Siegel, and Anjan V. Thakor, "Learning Corporate Control and Performance Requirements in Venture Capital Contracts," *International Economic Review* 31–2, May 1990, 365–381.

Competition in Banking, OECD 1989, 127 and 200–201.

"European Banking Integration in 1992," Salomon Brothers Stock Research, June 1989, 9.

Haller, Mark H., and John V. Alvitti, "Loansharking in American Cities: Historical Analysis of a Marginal Enterprise," *American Journal of Legal History* 21, 1977, 12–156.

Hellmann, Thomas F., and Manju Puri, "The Interaction Between Product Market and Financing Strategy: The Role of Venture Capital," *Review of Financial Studies*, Winter 2000, (13–4), 959–984.

Mayer, Martin, 1990, *The Greatest Ever Bank Robbery: The Collapse of the Savings and Loan Industry*, Charles Scribner's Sons, New York.

Mehr, Robert, *Fundamentals of Insurance*, 2nd Edition, Irwin, 1986.

Prudential Supervision in Banking, OECD 1987, Appendix X.

Reuter, Peter, and Jonathan Rubinstein, *Illegal Gambling in New York*: *A Case Study in the Operation, Structure and Regulation in An Illegal Market*, National Institute of Justice, 1982.

Vertanian, Thomas P., David H. Ansell, and Robert H. Ledig, "Federal Deposit Insurance Corporation Improvement Act of 1991 Prompt Corrective Action Matrix," a publication by the Financial Institutions Group at *Fried, Frank, Harris, Shriver & Jacobson*, 1993.

CHAPTER ♦ 3

The What, How, and Why of Financial Intermediaries

"All essential knowledge relates to existence, or only such knowledge as has an essential relationship to existence is essential knowledge."

Soren Kierkegaard: *Concluding Unscientific Postscript*

Glossary of Terms

Securitization: The act of converting an untraded (debt) claim, such as a bank loan, into a traded security by issuing claims against it and selling these claims to capital market investors. Essentially, securitization is a form of direct capital market financing with the bank acting as an originator and repackager of the loan.

Fractional Reserve Banking: A banking system in which banks must hold a specified fraction of their deposit liabilities as liquid assets.

Fiat Money: A form of money, the acceptance of which is mandated by law.

The Market Model: A model that states that the return on a security can be partitioned into a fixed component (called "alpha"), plus a component which is a multiple (called "beta") of the return on the "market" portfolio, plus a mean-zero residual term.

DIDMCA: The Depository Institutions Deregulation and Monetary Control Act passed in 1980. See Chapters 11 and 12 for details.

The Law of Large Numbers: Roughly speaking, a principle that says that if we have an infinitely large number of random variables in a sample, all of which are drawn from the same probability distribution, then the average realized value of the random variables in the sample will equal the statistical mean of the probability distribution from which they are drawn. Thus, if an individual divides his finite wealth equally across an infinitely large number of investments whose random payoffs are independent of each other, but are drawn from the same probability distribution, this individual's payoff from his investments will become (almost) *certain* and equal to the statistical mean of the probability distribution from which investment payoffs are drawn. A risk-averse individual would prefer to do this because it eliminates risk.

Event Study Methodology: A statistical approach commonly used in finance to evaluate the price impact of an event. The idea is to start with the assumption that the return on a stock can be described by the market model. Then, the next step is to estimate the values of alpha and beta by regressing the return on the stock against the return on the market for a sufficiently long time period prior to the event date and outside a 2- or 3-day time window around the event date. Given these estimated values, one can compute the average value of the residuals during the time window around the event date. If no new information was conveyed by the event, the average value of the residuals should be zero. If it is positive (negative), the event is interpreted as conveying good (bad) news.

Natural Monopoly: In some industries, due to economies of scale, the most economically efficient industry structure is to have only a single firm that is a natural monopoly.

Capital Requirements: The requirements that the bank keep a minimum amount of capital, consisting of equity, long-term debt, and other claims subordinated to deposits. See Chapters 2 and 11.

Portfolio Restrictions: Restrictions on the assets that banks can hold in their portfolios. See Chapters 2 and 11.

Introduction

As the following exchange between Levin and Sviyazhsky from Part III, Chapter 27 of Tolstoy's *Anna Karenina* indicates, most people know what banks and other financial intermediaries do.

> "Then what's your opinion? How should a farm be managed nowadays?"
> "What we have to do is to raise the standard of farming even higher."
> "Yes, if you can afford it! It's all very well for you, but... I'm not going to be able to buy any Percherons."
> "That's what banks are for."

As perceptive as this notion of banking is, we will need a deeper understanding of banks and other financial intermediaries in order to set the stage for the remaining chapters in this book. The simple view that banks exist to provide borrowing and lending services leaves us without answers to questions such as the following: (i) Why do we need *banks* to intermediate between borrowers and lenders, that is, why don't individual borrowers and lenders transact *directly* and avoid the cost of going through banks?[1]

1. A partial answer to this question was provided in Chapter 2.

(ii) What, if any, are the economies of scale in the production of financial services provided by banks, or, how large should banks be? (iii) Why do we regulate banks and other depository institutions so intrusively? (iv) If banks need to be regulated, *how* should they be regulated? (v) How should borrowers choose whether they should borrow from banks, or venture capitalists, or directly from the capital market?

To answer these and other questions, we need a framework that builds upon that provided in the previous chapter and illuminates the *essential* functions served by financial intermediaries. While we will not provide complete answers in this chapter to all of the questions posed above, our purpose is to provide a systematic way to think about these issues, so that we have a foundation for the discussions in subsequent chapters. The plan for this chapter is as follows. We begin with an anecdotal discussion of how a fractional reserve banking system arises from a simple goldsmith economy. After this informal discussion we provide a model of a bank that formalizes the goldsmith anecdote and helps us to understand the role of banks as well as the need to regulate them. These two sections provide answers to questions (i) and (iii) above, and a partial answer to question (iv). The next section introduces the fixed coefficient model as an extension of the goldsmith anecdote and examines its implications for monetary policy. The issue of economies of scale in the production of financial intermediation services is then taken up. This provides an answer to question (ii) above. Following this, we proceed to explain how banks can make nonbank contracting more efficient, and then we review empirical evidence in support of the view that banks are special. The ownership structure of depository institutions is analyzed next. We conclude with an examination of a borrower's choice of financing source to answer question (v) above.

Fractional Reserve Banking and the Goldsmith Anecdote

Fractional Reserve Banking

Chapter 2 explains what financial intermediaries do. We will now continue this discussion by examining how a rudimentary bank can evolve from a goldsmith, and how this leads to a theory of fractional reserve banking. What emerges too is a theory of bank regulation. According to this theory, regulation is an almost inevitable outgrowth of fractional reserve banking.

Modern banks produce *fiat* money on the basis of *fractional reserves*. These two facts account for much of the romance, mystique, and confusion surrounding finance. Laymen have difficulty understanding that money has value solely because of its universal acceptance as money.[2]

The fractional reserve aspect of banking is similarly vexing in that it seemingly involves sleight of hand. Fractional reserve banks fund themselves with liabilities that are convertible into cash on demand, but they hold only a fraction of such liabilities in the form of cash assets. Thus there is always some probability that withdrawals will exceed the available cash.

2. The acceptance of money is ultimately a social convention supported by the legal system, which recognizes money as an instrument for the legal discharge of debts. This view of money serves as the basis for arguing that seigniorage rightfully belongs to the community at large and should not be appropriable by private interests.

The evolution of monetary systems from commodity money—gold, silver, or whatever—to more abstract forms of money parallels the evolution of banking systems from warehouses, or 100 percent reserve banks, to modern fractional reserve banks. Both follow naturally from a collective desire to use scarce resources efficiently. However, these developments have side effects as well. The substitution of fiat for commodity money concentrates enormous economic power, for good or ill, in the hands of the monetary authority. Likewise, fractional reserve banking places enormous power in the hands of individual bankers, power to jeopardize the stability of the banking system in the pursuit of personal gain.

In what follows we shall explain the evolution of fractional reserve banking from its historical roots in warehousing. The explanation is stylized and anecdotal, and is meant to stress the natural aspects of the evolutionary process as well as the essential vulnerability of fractional reserve banking systems.

The Evolution of the Primitive Goldsmith Into a Bank

Think of a primitive setting in which gold is used as money—means of payment, or medium of exchange. By social convention, all debts are paid with gold and all purchases are made with gold. The system works well enough, but holding and transporting gold can be awkward. There is both a security problem and a convenience problem. The market response is to provide a warehousing service for gold. Hence the emergence of the goldsmith.

For a fee, the goldsmith provided secure storage facilities for gold. The owner of the gold would receive a warehouse receipt in exchange for her gold, with the understanding that the owner could present the receipt at her convenience to redeem the gold from the goldsmith.[3] The goldsmith's was a simple business. Like the furniture warehouse, the goldsmith provided safekeeping service for a fee. Simplicity itself!

Owners of gold gradually developed confidence in the goldsmith and gold flowed in and out of the goldsmith's coffers with tedious and profitable regularity. Whenever a gold owner wanted to make a purchase, she would travel to the goldsmith, withdraw the necessary gold and take it to the market. At the market, the gold would be exchanged for the desired goods and just as routinely, the seller of the goods would return the newly acquired gold to the goldsmith in exchange for a warehouse receipt.

As these trading and payment practices became more and more pervasive, and as the goldsmith's reliability became more and more established, repeated trips to the goldsmith were recognized as wasteful. Each time a purchase was desired, the buyer would need to run to the goldsmith for gold, only to have this trip repeated by the seller, who would return the gold from whence it came.

Ultimately, the warehouse receipt passes from the buyer to the seller, and the only purpose served by the two trips is to test the goldsmith's integrity. But as the goldsmith's reputation for integrity grows with time and experience, the need for these trips seems increasingly unnecessary. Gradually, trade is effected with the exchange of warehouse receipts and the gold remains undisturbed in the goldsmith's vault. But the willingness to accept warehouse receipts in lieu of gold rests on the belief that the gold is available on demand. Any suspicion of the goldsmith will undermine the use of

3. When transferable, it is the ownership of the receipt that governs the redemption.

the receipts as means of payment. But so long as the goldsmith can project confidence, there is a saving to be had by avoiding the trips to and from the goldsmith.

Seen from the vantage point of the goldsmith, the growing use of receipts as means of payment means smaller flows of gold into and out of the coffers. One can imagine a time series of data points that describe the gold holdings of the smithy through time. As the use of receipts gradually replaces gold, the goldsmith's gold inventory becomes less and less volatile. In the limit, as the receipts totally displace the gold, the goldsmith's inventory remains practically unchanged through time, unless newly mined gold flows into the system, or other extraordinary occurrences take place. It gradually dawns on the goldsmith that it is not really necessary to have a unit of gold for each outstanding receipt. This idea must have come as a revelation, an epiphany. To be sure, the strait-laced would recoil at the idea of issuing more receipts than one had gold, but if no one ever withdraws the gold, then what possible harm?[4] The naughty possibility of printing extra warehouse receipts changed the world. This discovery was the banking equivalent of the Newtonian Revolution, every bit as important to banking as gravity was to physics.

Instability of the Fractional Reserve Bank

The extra receipts could not be distinguished from their more authentic counterparts and they consequently served as means of payment as readily as did the authentic (those whose issue was occasioned by a deposit of gold) receipts. The extra receipts were loaned to borrowers and earned interest. Assume that these loans are illiquid, that is, they cannot be redeemed on demand, but rather must be held to maturity in order to realize their full value. This means that the goldsmith is providing a key *liquidity transformation* service by issuing liquid claims to depositors that are backed by illiquid loans to merchants. The pedestrian goldsmith was thus transformed from a warehouse clerk into a banker! To see this, consider the following before-and-after balance sheets.

Goldsmith (Before)		**Goldsmith (After)**	
Gold 100 oz.	Receipts 100 oz.	Gold 100 oz. Loan 10 oz.	Receipts 110 oz.

Notice that after the goldsmith crosses the Rubicon (becomes a banker), his liabilities of 110 ounces exceed his capability to satisfy them in the unlikely event that all receipt owners should seek to convert to gold simultaneously. This potential failure is because loans are *illiquid*.

Therefore, inherent in the lending is a potential catastrophe—insolvency of the goldsmith. Of course, if the receipt owners almost never withdraw their gold, the probability of insolvency is small, perhaps very small. However, and this is critical, the risk of ruin is endogenous. That is to say, the goldsmith chooses the probability of insolvency with his choice of how many extra receipts to print, or equivalently, with his choice of how many loans to make. Each extra receipt printed and loaned earns interest and so the temptation to print receipts is limited only by the goldsmith's concern for remaining solvent. He walks the knife-edge between avarice and anxiety.

4. In a rational expectations equilibrium, the gold owners would anticipate this behavior of the goldsmith and adapt (redeem randomly and sufficiently frequently) to avoid being exploited by the goldsmith. But for present purposes, let us ignore this.

Each extra receipt increases income, but at the same time increases the probability of insolvency; insolvency, of course, destroys the goldsmith's reputation and with it his ability to circulate and lend warehouse receipts.

Thus we see how the discovery of fractional reserve banking was a rite of passage, a loss of innocence. Notice, however, that conditional on the loans being repaid, the goldsmith holds assets equal to the value of his liabilities. Thus what we have here is a liquidity issue. The goldsmith can and will pay off all receipt holders, given adequate time and good loans. Nevertheless, the promise is to *pay on demand*, and this most assuredly cannot be done in *all* states of nature.

This is the essence of fractional reserve banking and its essential vulnerability. Such a system evolves quite naturally given maximizing behavior on the part of rational economic agents.

Regulation as a Stabilizing Influence

Left to its own devices, this kind of banking system is subject to periodic collapse. However, experience with fractional reserve banking eventually led to the discovery of a rather simple and straightforward remedy. Since the Achilles heel of the system is the illiquidity of the loans, bank runs could be averted if these assets could be liquefied. What was needed was a *bank for goldsmiths* that could lend against the collateral of a goldsmith's loans during those infrequent occasions of extraordinary redemptions. Indeed, in the 19^{th} century this was achieved in the U.S. through commercial bank clearing houses (CBCHs), which were private arrangements between banks that agreed to put their *combined* resources (the CBCH) behind each member in times of unanticipated liquidity drains. (See Chapter 9 for more on CBCHs.) Of course, such a bankers' bank would need virtually unlimited capacity, together with a commitment to the continuity of the system. The private arrangements did not possess such unlimited capacity, and this provided the rationale for a central bank to serve as a lender of last resort to the community of bankers. Since the central bank, which was typically government-owned, had the privilege of printing (or otherwise creating) money, the issue of limited capacity evaporated.

One more point deserves emphasis in connection with the evolution of a fractional reserve banking system with a central-bank-based lender-of-last-resort facility. Absent the central bank, there will always be a self-imposed limit on the volume of extra receipts printed. The fear of failure, loss of reputation, and the consequent inability to continue to lend warehouse receipts will discipline the inclination to expand lending indefinitely. Whatever this self-imposed limit, however, the introduction of the central bank acting as a lender of last resort will weaken the goldsmith's restraint. If the goldsmith knows that he can borrow against his otherwise illiquid loans, he will make more loans than if he could not use the loans as collateral. This is clear and obvious; and it is true even if the central bank charges a very high rate of interest for such emergency borrowings. Note that the interest rate for such loans is infinite in the absence of the central bank. Thus, the central bank introduces a kind of moral hazard, and this moral hazard is typically addressed by imposing cash asset reserve requirements that effectively limit the volume of a bank's lending on the basis of its cash assets. This is perhaps the most basic of prudential regulation. The point is that regulation is endogenous. It is responsive to a moral hazard arising from the introduction of the central bank as a lender-of-last-resort, which in turn is a response to a vulnerability inherent in fractional reserve banking. In turn, fractional reserve banking is a natural response to the transport costs and security concerns in a *laissez-faire* world of commodity money.

A Model of Banks and Regulation

That the very nature of banking necessitates regulation can also be seen in the perspective of a model in which money—rather than gold—is used as a medium of exchange. We will now develop in the box below a model that formalizes the anecdotal development of the previous section and also highlights some of the underlying informational assumptions in the analysis. The intuition is very similar to that in the earlier section.

The two-period model developed below is very simple.[1] It makes some assumptions that are not rigorously justified. Our intent is to give a broad-brush, intuitive treatment of how banks arise even in primitive economies and why it is necessary to regulate them. Before developing the model, we provide a summary of the notation used in Table 3.1

TABLE 3.1 The Notation

Notation	What it Means
y	Depositor's income in each period.
c	Depositor's consumption from income in each period.
s	Amount deposited in each period.
ϕ	Fee charged to depositor for safekeeping of deposits.
$\hat{\phi}$	Personal cost of safeguarding deposits.
α	Fraction of deposits withdrawn.
n	Number of depositors.
m	Number of merchants.
K	Merchant's cash flow.
K^*	High value of merchant's cash flow when K is random.
M	Loan to merchant.
p	Probability of theft.
r	Rate on return on bank's loan to a merchant.
b	Bank's cost of monitoring merchants.
u	Probability that $K = 0$, as assessed at date 0.
u_1	Value of u, as per updating at date 1.
u_h	High value of u_1.
U_ℓ	Low value of u_1.
L	Liquidation value of merchant's investment.
f	Amount depositors must spend to ensure that the bank safeguards and monitors.
j	Number of banks.

(*Continued*)

The Model: Consider an economy in which individuals are unsure of how safe their personal wealth is from theft. Thus, it pays to safeguard it. The individual can either safeguard it himself or he can pay someone else to do it. It is easy to imagine that not everybody is equally skilled in the art of safeguarding. So if you believe others are more skilled in safeguarding, you may wish to entrust safeguarding of your wealth to someone else, even though this involves paying a fee.[2] Since we will eventually reach this conclusion anyway, let us refer to you (the person who wishes to have his personal wealth safeguarded) as the depositor and the entity that safeguards your wealth as the bank. For now let us suppose there is only one depositor ($n = 1$) and only one bank ($j = 1$).

The depositor has an income of \$y in each period, of which \$c goes to personal consumption and $\$(y - c) = \$s$ goes to savings. These savings must be safeguarded. For now suppose there is nothing that the bank can do with this money except safeguard it. Let $\$\,\phi > 0$ be the fee that the bank charges to safeguard the depositor's savings. Safeguarding by the bank guarantees that the wealth will not be stolen. Also suppose that the depositor wishes to have his wealth safeguarded for only one period. Assuming that the discount rate is zero for everybody,[3] we see that the depositor's consumption at the start of the next period will be $s - \phi$ (his net saving in the first period) plus $\$y$ (his income in the second period). Since the depositor is paid $\$s - \phi$ for depositing \$s, the interest rate on his deposit is

$$(s - \phi - s)/s = -\phi/s < 0 \qquad [3.1]$$

If a negative interest rate surprises you, remember that our bank cannot make any loans and is providing the depositor a costly service. Assume for now that the bank must keep 100 percent reserves against deposits and that the depositor will fully withdraw at the end of the first period.

The Desirability of a 100 Percent-Reserves Bank: Suppose the probability of theft is p and it would cost the depositor $\hat{\phi} > \phi$ to safeguard his wealth to the extent that the probability of theft is eliminated. Thus, a necessary and sufficient condition for personal safeguarding to be optimal is that

$$\begin{gathered} s - \hat{\phi} > (1 - p)s \\ \text{or } \hat{\phi} < ps \\ \text{where } 0 < p < 1. \end{gathered} \qquad [3.2]$$

We will assume that (3.2) is satisfied. Clearly, since $\phi < \hat{\phi}$, the depositor will prefer to have the bank safeguard his wealth.

Note that in stipulating that the bank charges the depositor exactly what it costs the bank to safeguard, we have assumed that there is perfect competition[4] between banks that can all safeguard s at $\$\phi$. Suppose now that $n > 1$, so that there are possibly many depositors. It would be natural to assume that there are economies of scale in safeguarding, that is, it should cost less per dollar to safeguard \$ns as opposed to safeguarding \$s. For example, one armed guard may be able to safeguard \$100,000 just as easily as he can safeguard \$1,000. Indeed, if we were to assume that the cost of safeguarding \$ns is less than $\$n\phi$, the case for a large bank would be compelling, and we could even assume than $\hat{\phi} = \phi$, that is, no single individual is any more skilled than another in protecting wealth. But we will assume that there are no scale economies in safeguarding. In a sense, this makes our task harder, but it helps to reduce notation.

Suppose first that all the n depositors will surely withdraw at the end of the first period. In this case, it is easy to see that the interest rate will still be $-\phi/s$. A more interesting and natural case, however, is one in which *not* all depositors will withdraw at the end of the first period. Suppose a fraction α (where $0 < \alpha < 1$) of depositors will withdraw at $t = 1$ (the end of the first period) and the remaining fraction $1 - \alpha$ will withdraw at $t = 2$ (the end of the second period). For simplicity, we assume that α is known with certainty.[5] The sequence of events is described in Figure 3.1 below.

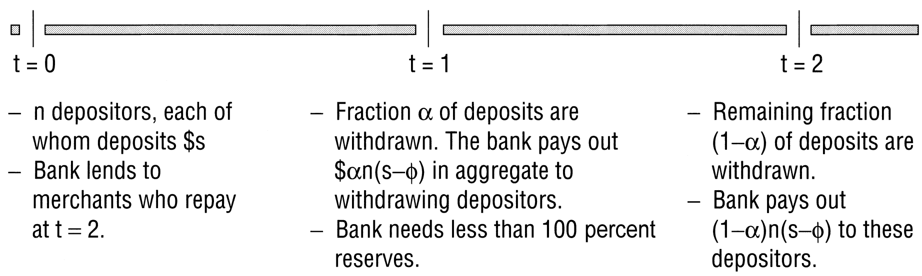


FIGURE 3.1 Sequence of Events

A Bank That Borrows and Lends: If the bank cannot invest any of the deposits it receives, then funds will lie idly in the bank. Note, though, that there is an opportunity to invest in this case.[6] At $t = 1$, the bank only needs to have $\$\alpha n(s - \phi)$ to meet deposit withdrawals. Suppose now that it is possible for the bank to make investments at $t = 0$, but that these investments will pay off only at $t = 2$. Let r be the rate of return to the bank on these investments.[7] We can imagine that the investments are loans to merchants who want to finance the setting up of shops, but do not have any funds of their own. Each merchant needs \$M, where $M > s$, so that if the merchant were to borrow directly from depositors, he would need to approach more than one depositor (in fact, he would need to approach *M/s* depositors). Further, there is a *moral hazard* problem in dealing with the merchant in that he has a preference for absconding with the \$*M* he borrows rather than setting up a shop. If his actions are not monitored, he will abscond and the lender will not be paid back at all. However, at a cost of \$b it is possible to monitor the merchant so that he indeed puts his borrowed funds to the stated use of setting up a shop that will generate some cash flow of $\$K > M(1 + r)$ at $t = 2$. As a start, let us suppose that \$K is a sure cash flow. We will introduce uncertainty shortly.

First consider the merchant's problem if he approaches M/s depositors directly. His net expected payoff will be

$$K - M(1 + \mathrm{r}) - (\mathrm{b} \bullet M/s) \qquad [3.3]$$

since, in addition to interest, he will be charged for monitoring. Each depositor will have to individually monitor the merchant since none can rely on his cohorts to do so.[8] Now, if the merchant approaches a bank, which in turn acquires \$M in deposits from *M/s* depositors, we will have a different outcome. The bank's monitoring cost

(*Continued*)

will be \$b. If the bank charges the merchant exactly what it costs the bank to monitor, then the expected payoff to the merchant will be

$$K - M(1 + r) - b. \qquad [3.4]$$

Comparing (3.3) and (3.4) we see that the merchant is clearly better off going to the bank.

Since the merchant pays the bank only at $t = 2$, the bank will have to make sure that it will have enough money at $t = 1$ to pay off depositors who withdraw then. Suppose there are m merchants (borrowers) and n depositors. Then, the bank loans out \$mM and takes in \$ns in deposits. Let $ns > mM$ (this will be shown to be necessary in a moment). Since \$mM are loaned out, the bank doesn't need to worry about safeguarding that money from outright theft (it just needs to monitor the merchants it lends to). Thus, \$(ns – mM) must be safeguarded. The safeguarding cost is $(ns - mM)\phi/s$, since it costs ϕ/s to safeguard \$1.

Hence, the bank promises to pay depositors

$$ns - (ns - mM)\phi/s \qquad [3.5]$$

in the aggregate if it does not pass along to the depositors any of its profits from lending to merchants. Since a fraction α of deposits are withdrawn at $t = 1$, those depositors get $\alpha[ns - (ns - mM)\phi/s]$, which you will notice is more (by an amount $(mM\phi/s)$) than what these depositors received previously. That is, the fact that part of the money is being loaned out instead of being kept in the bank's vault itself economizes on safeguarding costs. Although the loaned money must be monitored, these monitoring costs are paid by borrowers, so that depositors realize a saving in safeguarding costs.

To ensure that the bank will have sufficient funds to meet deposit withdrawals at $t = 1$, it must choose m to satisfy

$$\alpha\,[ns - (ns - mM)\phi/s] = ns - mM - (ns - mM)\phi/s - mb. \qquad [3.6]$$

To understand (3.6), note that the left-hand side is the amount the bank must pay out to those depositors who withdraw funds at $t = 1$. On the right-hand side, $ns-mM$ is the amount of money the bank has left over in reserves after it is through lending to the m merchants. From this it must spend an amount $(ns-mM)\ \phi/s$ to safeguard its reserves and an amount mb to monitor the m merchants.[9] Solving (3.6), we get

$$m = (1 - \alpha)ns(s - \phi)/\{M[s - (1 - \alpha] + bs\} \qquad [3.7]$$

Thus, as long as the bank lends to exactly as many borrowers as stipulated in (3.7), there will be no risk of withdrawals exceeding the bank's available cash reserves at $t = 1$.

Note now that the bank makes an aggregate net profit of mMr on its lending activities. This is because it is being compensated exactly for its monitoring cost by borrowers, and its safeguarding cost by deposit interest rate, although higher than $-\phi/s$ (as in the previous case when all deposits were idle), is still negative. This positive profit will attract entry by competing banks, and the resulting competition for depositors' funds will drive up the deposit interest rate. In a competitive equilibrium, each bank will earn zero profit. This will happen when the bank's profit of mMr is divided equally among the n depositors, so that each depositor gets

$$\frac{ns - [(ns - mM)\phi/s] + mMr}{ns}$$

per dollar of deposits. Thus, the deposit interest rate is now

$$\frac{ns - [(ns - mM)\phi/s] + mMr}{ns} - 1$$
$$= \frac{mMr - [(ns - mM)\phi/s]}{ns} \qquad [3.8]$$

If we assume that r is high enough to ensure that the numerator in (3.8) is positive, then the depositors get a positive rate of interest on their deposits.

We have taken you through a sequence of steps to show how a bank, like the goldsmith in the previous section, can develop from a simple caretaker of other people's wealth into an institution that borrows and lends money. As you must have noted, informational problems play a key role in bringing our bank to life. Banks solve two types of moral hazard problems in our simple world. First, they help to cope more efficiently with the "social" moral hazard problem of theft. Second, they also help to cope more efficiently with moral hazard in lending, which, as you know from Chapter 1, is a type of agency problem.

Do We Need to Regulate This Bank?: So far, however, there has been no need for a regulator. But that is simply because we have made numerous strong assumptions. One of them is that it is possible to monitor merchants so efficiently that they'll always repay their debts fully if they are monitored. In reality, merchants may sometimes have poor cash flows even if they do their best. That is, suppose that, viewed at $t = 0$, their cash flow K is a random variable that is 0 with probability u and K^* with probability $1 - u$. We'll assume that setting up a shop is a positive net present value (NPV) exercise for the merchant, so that

$$(1 - u)K^* > M(1 + r). \qquad [3.9]$$

Suppose that this in itself does not affect the behavior of depositors in terms of their withdrawal policies. But at $t = 1$, depositors may learn something more about the likelihood that merchants may fail. For simplicity, assume for now that merchants have perfectly correlated prospects, so that they all either fail ($K = 0$) or succeed ($K = K^*$). Let us refer to the updated probability of failure that depositors assess at $t = 1$ as u_1. If there is good news, $u_1 < u$ (the probability of failure they assessed at $t = 0$) and if there is bad news, $u_1 > u$. We can think of u as the expected value of u_1 assessed by depositors at $t = 0$. Suppose u_1 can take one of two values: $u_1 = u_h$ for bad news and $u_1 = u_\ell$ for good news, where $u_h > u_\ell$. Suppose that those depositors who intended to withdraw at $t = 2$ will in fact change their minds and withdraw at

(*Continued*)

$t = 1$ if they get bad news[10], that is, if $u_1 = u_h$. If they get good news, they'll withdraw at $t = 2$.

The bank now faces a problem. If depositors get bad news, all depositors withdraw at $t = 1$. The bank will have insufficient funds to meet withdrawals (unless it keeps 100 percent reserves and does not lend to any merchants). Suppose that in this case the bank is empowered to call back all of its loans prematurely and this forces merchants to liquidate their businesses prematurely. Let L be the liquidation value of the merchant's shop at $t = 1$ (which, for simplicity, is independent of the information received by depositors at $t = 1$). Assume L is a very small number (much smaller than K^*). So, if all depositors wish to withdraw funds at $t = 1$, and if the bank proceeds to lend exactly the same amount at $t = 0$ as it did in the previous case, then there will only be $\$\alpha[ns - (ns - mM)\phi/s] + mL$ to pay depositors. Moreover, the premature liquidation of merchants' shops will be socially inefficient if L is so small that $L < (1 - u_h)K^*$. This is similar to the illiquidity problem of the goldsmith.

There is no way that the bank can prevent this unless it keeps all of its deposit funds idle, in which case it doesn't matter when depositors withdraw. However, this would *not* be fractional reserve banking; it would hardly be a bank as we know it. This is where a government regulator can help. Suppose it agrees to insure all deposits for the full promised payment by each bank. Then we see that those depositors who originally planned to withdraw at $t = 2$ have no reason to change their minds since the value of u_1 is now irrelevant to them; the deposit insurer has made their claims risk free! That is, this form of regulation makes banking viable when it otherwise could not have been.

This seems to be a wonderful solution and it definitely has its merits. But lest we get carried away with its virtues, let's pause and complicate things a bit more. Since banks are competitive and earn zero profits, they may wish to underspend on either safeguarding or on monitoring borrowers. Once the terms of their loan and deposit contracts are set, they could profit from spending less on safeguarding and monitoring than originally promised. Depositors will rationally anticipate this moral hazard and try to prevent it. Suppose that each depositor could spend a small amount of money, say \$f, to make sure that the bank expends the promised resources on safeguarding and monitoring. We can show, given appropriate assumptions, that depositors will find it in their own best interest to do so.

1. This model has some features found in Millon (1983). Other papers dealing with the existence of financial intermediaries are Leland and Pyle (1977), Campbell and Kracaw (1980), Diamond (1984), Ramakrishnan and Thakor (1984), Millon and Thakor (1988), Boyd and Prescott (1986), and Allen (1990).

2. Naturally, this fee should be less than what it would cost you to safeguard your own wealth with the same efficacy.

3. This is a harmless assumption and can be easily dropped without affecting this analysis.

4. For those of you well-versed in different notions of competition in economies, we have in mind Bertrand competition here.

5. We will discuss later what happens if α is random.

6. Actually, even in the previous case in which all deposits are withdrawn at $t = 1$, the bank could invest at $t = 0$ in assets that pay off at $t = 1$.

7. We will not go into the details of how r is determined.

8. It is obvious that we cannot have an equilibrium in which no depositors monitor, because then it pays for at least one to monitor. To justify an individual depositor's decision to monitor, we must assume that there is some uncertainty that some depositors will not monitor (otherwise, every depositor will wish to "free ride" on the

Summary: Thus, one way to prevent bank runs and instability is for the government to provide deposit insurance, which is an alternative to the lender-of-last-resort (discount window) facility provided by the regulator in our earlier goldsmith example. But there is a fly in this ointment. When there is deposit insurance, why should any depositor care about whether the bank safeguards and monitors with the requisite vigilance? Each depositor's payoff is guaranteed and independent of the bank's actions. Hence, none will find it personally profitable to spend anything on watching over the bank to ensure that the bank expends the promised resources in safeguarding and monitoring the merchants it lends to. In other words, deposit insurance weakens or even destroys the private market discipline imposed on banks. The burden of keeping the bank in check shifts now from the market to the regulator. To achieve its objective, the regulator will have to come up with ways to dissuade the bank from exploiting the deposit insurance umbrella. In other words, the moral hazard engendered by one form of regulation, namely deposit insurance, creates the need for other forms of regulation (such as capital requirements, portfolio restrictions, and so on).

We have now completed the story we set out to tell in this section. Regulation is not just the outcome of some political agenda. It arises quite naturally from the very forces that give rise to banks. Once regulation arises to instill public confidence in banking and make banks viable entities, it creates its own moral hazards that necessitate further regulation.

The Macroeconomic Implications of Fractional Reserve Banking: The Fixed Coefficient Model

In this section we examine the implications of fractional reserve banking for monetary policy. The discussion developed here formalizes some of the macroeconomic implications of the goldsmith anecdote presented earlier.

monitoring of his cohorts). One way to do this is to assume that *each* depositor believes that there is a random fraction θ of the remaining $(M/s) - 1$ depositors who are simply incapable of monitoring, but no one (except those incapable depositors themselves) can identify these depositors. Thus, each of the depositors will still charge for monitoring but will not spend \$b. Suppose θ can be 0 with probability q_0 and 1 with probability $1 - q_0$ (when $\theta = 1$, each depositor who can monitor believes that he is pivotal in that no one else will monitor). Then, if a depositor who monitors chooses not to do so, his expected payoff will be (he always assumes that all other depositors capable of monitoring will indeed monitor) $b + q_0 s(1 + r) - s = b - (1 - q_0)s + q_0 sr$. And if he chooses to monitor, his expected payoff will be $s(1 + r) - s = sr$. Thus, it is a (Nash) equilibrium to monitor if $sr > b - (1 - q_0)s + q_0 sr$ or if $b < (1 - q_0)s(1 + r)$. Thus, if the uncertainty about incapable depositors is sufficiently large in the mind of each capable depositor (that is, $1 - q_0$ is sufficiently high) and if the monitoring cost b is low relative to the payoff $s(1 + r)$ from successful monitoring, each capable depositor will monitor in a Nash equilibrium.

9. We are assuming here that safeguarding costs are paid just after $t = 0$ and monitoring costs are paid just before $t = 1$. Note that since the merchants repay the bank only at $t = 2$ and monitoring must proceed at $t = 1$, the bank must initially pay the necessary monitoring costs and then recover these costs from borrowers at $t = 2$ through a loan interest rate that is grossed up to reflect this cost.

10. Let us not worry about why they might wish to do this. We want to give you an idea of the underlying concepts without being too rigorous. It is possible to make these ideas work more rigorously.

The Fixed Coefficient Model

The Fixed Coefficient Model (FCM) is the standard textbook description of the banking firm and industry; it emphasizes the asset-transformation function of financial intermediaries. The bank's effort to maximize its profit is captured only implicitly. Consider a bank's balance sheet

Bank Balance Sheet

R	D
M	E

where R is the reserves of the bank comprised of deposits held at the central bank, M is the bank's earning assets (loans to merchants), D is the bank's deposit liability (think of this as $n \times s$ in the context of the model in the previous section), and E is the bank's equity. We can now write the balance sheet identity for the bank as:

$$R + M = D + E. \qquad [3.10]$$

Moreover,

$$R = rD, \text{ with } 0 < r \leq 1. \qquad [3.11]$$

Equation (3.11) represents the fact that banks hold cash or liquid asset reserves proportional to deposits in order to insure against deposit withdrawals and/or to satisfy legal reserve requirements. The fixed coefficient, r, can be interpreted either as a legal reserve requirement or a voluntary behavioral parameter (that is, reserves that the bank chooses to voluntarily hold). Actually, it should be interpreted as the greater of the two. In any case, the parameter relates to *liquidity or withdrawal risk*. That is, it is the bank's safeguard against a fraction (α in the context of the model in the previous section) of deposits being unexpectedly withdrawn. Next, we have

$$E = eL, \text{ with } 0 < e \leq 1. \qquad [3.12]$$

Equation (3.12) represents the fact that banks hold capital reserves in some fixed proportion, e, to loans in order to protect against *insolvency or default risk*. The parameter e can be interpreted as a regulatory capital requirement and/or a voluntary behavioral parameter, or, more accurately, the greater of the two.

An Illustration of the FCM

Let us now consider the FCM in a (competitive) banking industry with zero equity ($e = 0$) where banks have only two assets (reserves held in the form of deposits at the Federal Reserve and loans to the public) and one liability (customer deposits). We shall further assume a 20 percent effective legal reserve requirement ($r = 0.2$). The assumption that $e = 0$ is an extreme representation of the assumption that the capital requirement is not binding.

Now suppose Bank A receives a $1,000 deposit.

Bank A

Required Reserves 200	1000 Deposits
Excess Reserves 800	
Total Reserves 1000	

Since it has excess reserves of $800 and since it earns nothing on either its required reserves or excess reserves, the bank seeks to eliminate its excess reserves by making a loan of $800:

Bank A

Total Reserves 1000	1000 Deposits
Loan 800	800 Deposit

The funds loaned by Bank A, although possibly initially deposited with Bank A, are soon withdrawn and deposited in another bank, say Bank B. This leaves Bank A with

Bank A

Required Reserves = Total Reserves 200	1000 Deposits
Loans 800	

But Bank B has

Bank B

Required Reserves 160	800 Deposits
Excess Reserves 640	

and Bank B now lends away its excess reserves, so that:

Bank C

Required Reserves 128	640 Deposits
Excess Reserves 512	

The $640 loaned by Bank B is now deposited in Bank C. The process continues *ad infinitum*. At the Federal Reserve, the initial deposit would be a credit of $1,000 to Bank A.

Federal Reserve

	1000 Deposit A

What is the offsetting asset (liability) entry?

When the \$800 is withdrawn from A and deposited in B, the Federal Reserve would show

Federal Reserve

	200 A 800 B

Notice that the original reserve creation (the \$1,000 deposit received by Bank A) spurred deposit expansion, and the deposit expansion redistributes the reserves across the banking system. However, the deposit expansion does not affect the level of reserves in the banking system. In fact, deposit expansion absorbs reserves. What this illustration of the FCM shows is that the bank's incentive to hold reserves—either voluntarily to protect against unanticipated deposit withdrawals or to satisfy a regulatory reserve requirement necessitated by the moral hazard created by the lender-of-last-resort facility—results in less lending than would be possible without reserve requirements. Moreover, it also affects the redistribution of liquidity throughout the entire banking system. This has macroeconomic implications that we explore below.

The FCM and Monetary Policy

The FCM helps us to understand the basic elements of how monetary policy works. There are three major tools of monetary policy: (i) open market operations, (ii) reserve requirement changes, and (iii) discount rate changes. These three tools are used in varying degrees to influence the stock of money and interest rates.

Open market operations are sales and purchases of government securities (Treasuries) by a special committee of the Federal Reserve. These sales and purchases affect the amount of reserves available to banks and thus, as indicated in previous subsections, the amount of lending. To see this, suppose the Fed buys \$1,000 in Treasury securities from the nonbank public. Then the nonbank public's balance sheet will be

Public

Bonds – \$1,000 Deposits of cash in Bank A + \$1,000	Liabilities unchanged

and Bank A's balance sheet will be

Bank A

Required Reserves 200 Excess Reserves 800	1000 Deposits

The $800 is now available to Bank A for lending. This means that the initial open market operation of purchasing Treasuries leads to an increase in lending by banks. Another way to view this is that the government has reduced public debt (by buying back government securities) and facilitated an increase in private credit. The open market operation of selling government securities has the opposite effect.

It is obvious that a change in reserve requirements will also affect bank lending. Any increase in reserve requirements will reduce the amount of deposits available for lending, and any reduction in reserve requirements will increase the amount of deposits available for lending. Thus, when the Federal Reserve desires to implement a contractionary monetary policy (to cool down inflation, for example), it can raise reserve requirements; similarly it can lower reserve requirements when it wishes to stimulate the economy.

Finally, the discount rate, which is the rate charged by the Fed to member banks for short-term borrowings from the Federal Reserve, also affects monetary expansion/contraction. By raising the discount rate, the Fed makes it more costly for banks to borrow and build up reserves, and therefore effectively reduces the reserves available to banks. This reduces lending. Likewise, a lowering of the discount rate facilitates increased lending.

This analysis is predicated on the "classical" assumption that the binding constraint on bank lending is the reserve requirements. If the capital requirements e [recall equation (3.12)] were binding instead, the effects of monetary policy can be very different indeed, as we will see in Chapter 10.

Large Financial Intermediaries

The theories from which we borrowed some of the ideas in the previous section suggest that financial intermediaries should be very large. These arguments are based on diversification. They explain why banks should be large. Similar intuition applies to nondepository financial intermediaries as well. In this section we develop this argument. We focus on the basic intuition; the mathematics can be found in Appendix 3.1. It leads to a rationale for *nondepository* financial intermediaries like investment banks, Standard & Poor's Value Line, credit rating agencies, financial newspapers, Moody's check guarantee services, portfolio managers, econometric modelers, consultants, and accounting firms.

What the theoretical research has shown is that F.I.s are optimally infinitely large regardless of whether they are brokers or asset transformers. That is, an F.I. is a "natural monopoly." We explain why below.

Brokerage as a Natural Monopoly: Consider a broker that specializes as an information producer. One problem that the broker's customers must be concerned about is that of information reliability. This is a key issue in information production. How do these customers know that the information the broker provides is accurate and reliable? One possible way to determine this is for customers to noisily assess the reliability of the information provided by the broker, and compensating the broker more when information is judged to be more reliable. This can be done either via reputational mechanisms – attaching higher reputation for reliability to a broker whose past information has turned out to be higher quality – or by comparing the broker's information to that available from other sources.

Now, if we are dealing with a single information producer, it can be quite costly to ensure that he will use reliable information, even if we can have a noisy assessment of this reliability. This becomes a little less costly if we are dealing with a producer who is a member of a *team* of information producers because then, by producing reliable information, *each* producer benefits not only himself (by making it more likely that he will obtain higher compensation) but also the team, and a share of the team's benefits accrues to each individual producer. This is an effective mechanism as long as the team members can monitor each other to ensure that nobody gets a "free ride." As the size of the team grows, more and more independent payoffs of individual producers are being pooled together before being divided equally among the team members, so that the resulting diversification reduces the risk in each member's compensation. The risk-averse information producers are thus made better off and they demand less compensation on an expected value basis to produce information. This makes the buyers of information better off. And the benefit keeps growing as the broker gets larger. That is, brokerage is a natural monopoly.

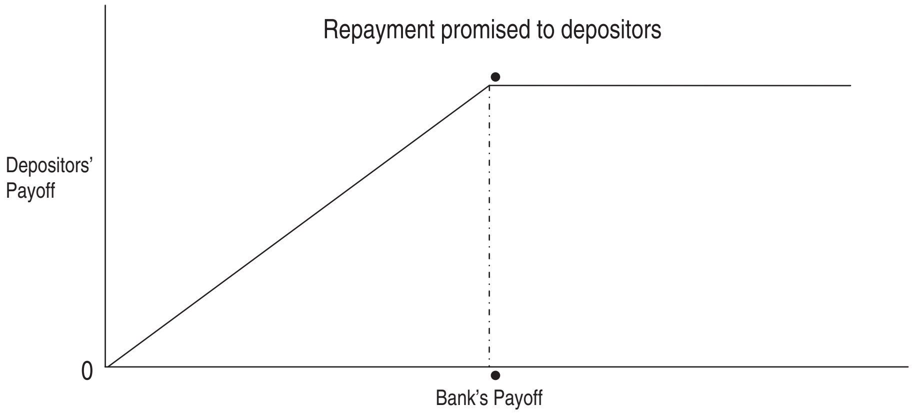
Another economic benefit from growing large comes from *information reusability*, which was discussed in Chapter 2. When information is cross-sectionally reusable, the larger the number of information producers in the intermediary, the greater is the benefit of information reusability. The reason is that information can be reused by a greater number of information producers within the intermediary, and yet the cost of acquiring information needs to be incurred only once.

A strong implication of this analysis is that investment banks, financial newsletters, credit-rating agencies, and other information producers can benefit from growing large. A caveat is that individual members can continue to monitor (and trust) each others as the organization grows large. If not, "free rider" problems will crop up, and it may not be beneficial to grow beyond a certain size because of the difficulty of implementing effective internal controls.

Asset Transformation as a Natural Monopoly: Now consider an asset transformer like a bank. It borrows money from depositors and makes loans. Its advantage in being large comes from two sources.[5] First, suppose multiple depositors are needed to finance a single bank borrower and the borrower's creditworthiness has to be established through costly credit analysis. Then having a bank perform this credit analysis once conserves screening resources compared to a situation in which all the depositors engage in costly screening of the borrower. That is, a bank eliminates duplicated screening. Second, the depositors' payoff is a debt contract, it is a concave function of the bank's payoff as shown on the next page.

Because the depositors' payoff is concave, they behave as if they are risk averse. Hence, they can be made better off by reducing the risk they face, and the benefit of this is a lower interest rate on deposits. The bank can do this by diversifying its risk across many different borrowers. And, because the benefit of diversification keeps growing with size, the bank is a natural monopoly.

5. The discussion below is based on a model developed by Diamond (1984).



How Banks Can Help to Make Nonbank Financial Contracting More Efficient

We have spent quite some time examining the flow of services that banks and other F.I.s produce. These services essentially take the form of intermediating in different ways between the users and providers of capital and of reducing their costs of exchanging capital. It has been suggested that banks not only permit the capital that flows through them to be exchanged at lower cost, but they also lower the cost of capital exchange between other parties.[6]

To understand this argument, let us examine the role of bank loans in a borrowing organization's information process. It is worthwhile to draw a distinction between *inside* and *outside debt*. Inside debt is defined as a contract in which the creditor has access to information about the borrower not otherwise publicly available. The creditor may even participate in the borrower's decision process. This could be achieved, for example, by the creditor having representation on the borrower's board of directors. Bank loans are inside debt. By contrast, outside debt is defined as publicly traded debt in which the creditor depends on information about the borrower that is publicly available. Commercial paper and publicly traded corporate bonds are examples of outside debt.

Bank loans offer a special advantage in this regard. They are usually of short maturities. This means they must be periodically renewed. These renewals are accompanied by bank evaluation of the borrower's ability to meet fixed payment obligations. Thus, if the bank renews a borrower's loan, it sends a positive signal about the firm to its other creditors. Note that credibility of this signal derives from the fact that the bank "puts its money where its mouth is" when it renews the loan. Given this credible and positive signal, other higher-priority creditors find it unnecessary to expend their own resources to duplicate the bank's evaluation. Thus, bank loans help to reduce duplication in borrower evaluation by multiple creditors.[7]

Banks also have a cost advantage in making loans to depositors.[8] The ongoing history of a borrower as a depositor communicates valuable information to the bank about the borrower's cash management activities. This permits the bank to assess the

6. See Fama (1980).

7. The argument that banks can lower the contracting costs of other parties can also be found in Fama (1990). For empirical work that follows upon the study discussed in this section, see Lummer and McConnell (1989).

8. This has been suggested, for example, by Black (1975) and Fama (1980).

risks of loans to depositors and to monitor these loans at lower cost than other (competing) lenders. This consideration is particularly important in short-term loans that are rolled over because of the relatively more frequent borrower assessments. This hypothesis has empirical validity in the observation that most short-term debt is in the form of bank loans.

The Empirical Evidence: Banks Are Special

It turns out that there is some interesting empirical support for the theories we have presented thus far. The central question of empirical interest to us is whether bank loans are unique, that is, do they provide any special service with their lending activity that is not available from other lenders? To answer this question we can examine the stock price responses to announcements of bank loans and other types of debt such as private placements of debt and public debt issues. The empirical evidence is that there is a positive and statistically significant stock price response to a borrower's acquisition of a bank loan. Further, the positive market reaction is not common to *all* private debt placements. There is, for example, a negative stock price response to debt placed privately with insurance companies. These findings seem to suggest that bank loans are unique.[9]

To examine these results let us first look at Table 3.2, which gives the distribution of announcements of different types of debt contracts for NYSE and AMEX firms.

Although there is no noticeable pattern in bank loans through time, there are two interesting observations. First, privately placed debt has been declining through time. Second, among all privately placed debt (bank loans plus other privately placed debt), bank loans dominate to the tune of 68.38 percent.

TABLE 3.2 Distributions by Year of Announcements of Bank Credit Agreements, Privately Placed Debt, and Publicly Placed Straight Debt for a Random Sample of 300 NYSE and AMEX-Traded Nonfinancial Firms for the Period 1974–1983

Year of Announcement	Bank Loan Agreements	Privately Placed Debt	Public Straight Debt
1974	9	4	5
1975	11	7	13
1976	7	7	8
1977	8	7	4
1978	1	8	6
1979	8	1	9
1980	11	1	10
1981	9	1	9
1982	10	1	16
1983	6	0	10
Total	80	37	90

Source: James, C., "Some Evidence on the Uniqueness of Bank Loans," *Journal of Financial Economics* 19, 1987, 217–235.

9. See James (1987).

In Table 3.3 we provide descriptive statistics for different types of debt.

As this table shows, firms using private placements and bank loans are on average smaller than firms using public offerings of debt. The average firm size in both the bank loan sample and the private placement sample is about 25 percent of the average firm size in the public debt sample. This evidence is consistent with the theory discussed thus far. Problems of moral hazard and particularly of asymmetric information can be expected to be more severe for smaller, lesser-known firms. Hence, banks have a greater relative contribution to make in resolving these problems in such firms. Not surprisingly then, we find that bank loans are the dominant source of debt financing for small firms.

Let us now see how the stock prices of borrowing firms react to the announcements of various forms of debt. This evidence is presented in Table 3.4.

The abnormal stock return here is defined in the usual fashion as the deviation of the realized rate of return from the expected rate of return given by the market model. That is, the abnormal stock return for firm j over day t is defined as

$$R_{jt} - \left(\hat{\alpha}_j + \hat{\beta}_j R_{mt}\right)$$

where R_{jt} is the rate of return of security j over day t, R_{mt} is the rate of return on the market portfolio over the same period, and $\hat{\alpha}_j$ and $\hat{\beta}_j$ are the ordinary least squares estimates of the market model parameters for firm j.

The average abnormal stock return for bank loan agreements in Table 3.4 is positive and statistically significant at the 0.01 level. In addition, two-thirds of the abnormal stock returns are positive. The negative average abnormal stock return associated with the announcement of a public offering of debt is not statistically significant.

If the positive response to bank loan agreements results from some benefit of inside debt not unique to banks, then one would expect to observe a similar response to debt that is privately placed with insurance companies. However, as Table 3.4 indicates, the response to the announcement of privately placed debt is −0.91 percent, which is statistically significant at the 0.10 level. Moreover, the difference between the average abnormal stock returns of bank loan agreements and privately placed debt is statistically significant at the 0.01 level.

TABLE 3.3 Descriptive Statistics for Commercial Bank Loans, Privately Placed Debt, and Publicly Placed Straight Debt for a Random Sample of 300 NYSE and AMEX-Traded Nonfinancial Firms for the Period 1974–1983

	Type of Borrowing					
	Commercial Bank Loans (Sample Size 80)		*Privately Placed Debt (Sample Size 37)*		*Public Straight Debt (Sample Size 90)*	
Descriptive Measure	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**
Debt amount (millions of dollars)	72.0	35.0	32.3	25.0	106.2	75.0
Firm size (millions of dollars)	675	212	630	147	2506	1310
Debt amount/market value of common stock	0.72	0.46	0.52	0.25	0.26	0.15
Maturity of debt	5.6	6.0	15.34	15.0	17.96	20.0

Source: James, C., "Some Evidence on the Uniqueness of Bank Loans," *Journal of Financial Economics* 19, 1987, 217–235.

TABLE 3.4 Average Two-Day Percentage Abnormal Stock Returns on the Announcement of Commercial Bank Loans, Privately Placed Debt, and Publicly Placed Straight Debt Offerings for a Random Sample of 300 NYSE and AMEX-Traded Nonfinancial Firms for the Period 1974 to 1983

Type of Event	**Abnormal Stock Returns**	**Proportion Negative (Sample Size)**
Bank loan agreement	1.93%	0.34 (80)
Privately placed debt	−0.91%	0.56 (37)
Public straight debt	−0.11%	0.56 (90)

Source: James, C., "Some Evidence on the Uniqueness of Bank Loans," *Journal of Financial Economics* 19, 1987, 217–235.

It is possible that the differences in abnormal stock returns across different types of debt agreements could be due to systematic differences in maturity and purpose of borrowing, that is, the data may not indicate anything special about bank loans *per se*. To check this possibility, we would like to know the share price responses to the announcements of bank loans, private placements, and public debt offerings, all with the same characteristics. The evidence on this score suggests that differences in abnormal performance across these different sources of borrowing are not solely due to differences in the characteristics of the loan or differences in the characteristics of borrowers (such as size, for example). That is, the results are robust. The overall conclusion to be reached from this empirical evidence is that banks are special.

Ownership Structure of Depository Financial Institutions

Depository institutions have two types of ownership forms: stocks and mutuals. Agency theory predicts that ownership form has a significant effect on the incentives and the operating efficiency of the firm. In this section, we will review the theoretical bases for this prediction and also look at some empirical evidence.

Commercial banks are exclusively stockholder-owned. Mutuals are common among insurance firms, MSBs (mutual saving banks), and S&Ls (savings and loan associations), although many mutual S&Ls have converted into stockholder-owned organizations in recent years. We will proceed as follows. First, we will examine how mutuality affects the resolution of agency and other problems. Then, we will seek an explanation for why S&Ls were dominantly mutuals and why the recent wave of conversions to stock ownership. Finally, we will review some relevant empirical evidence.

Mutual Versus Stocks

The residual claimants in a mutual are customers. These are the policyholders of mutual life insurance companies, the depositors of MSBs, and the depositors of mutual S&Ls. For purposes of this discussion, we will limit ourselves to mutual S&Ls.

There are two key differences between a stock and mutual S&Ls. First, the owners of a stock S&L are its stockholders, whereas the owners of a mutual S&L are its depositors (and possibly its borrowers). Second, a stock S&L can increase its capital by selling common stock, whereas a mutual S&L cannot.

Consider the first difference. In a stock S&L, shareholders have a well-defined ownership right, which implies: (i) a claim to residual profits, (ii) a right to vote for the board of directors and change control of the organization, and (iii) a right to dissolve the organization. On the other hand, in a mutual S&L, the ownership rights of depositors are much weaker. As for (i), depositors in a mutual are much more like creditors than shareholders since they cannot force the mutual to pay them more than the promised interest and principal on their claims. Although in principle depositors have ownership claims to the mutual's current earnings, these claims are not transferable, and the earnings can be retained indefinitely by the institution as net worth.[10] As for (ii), while mutual S&L depositors have voting rights, these are quite limited and are often signed over to management at the time of opening of accounts.[11] Finally, as for (iii), even though a depositor can withdraw his deposits and thereby partially liquidate the mutual fund,[12] depositors have had little incentive to do so because of deposit insurance, especially when interest rate ceilings bounded the return to depositors.[13]

Thus, it is imperative to distinguish between *de jure* and *de facto* ownerships in a mutual. The *de jure* ownership (legal ownership) rests with the mutual's customers. It is, however, largely vacuous. The *de facto* ownership [control of (i), (ii), and (iii)] rests with the managers and the government (which provides deposit insurance).

Of course, the inability of owners to completely control the institution—and the resulting agency problem—is encountered in stockholder-owned institutions as well. Both stock and mutual S&Ls are administered by managers whose goals may differ from the goals of the owners. However, the two types of S&Ls differ with regard to the ability of the owners to monitor managers. Stockholders have greater control over the activities of managers because control can be consolidated through the purchase of stock.

Agency Problems in Stocks and Mutuals

The above discussion suggests that agency problems in mutuals should be greater than those in stockholder-owned institutions. There are two ways in which we can measure the incidence of agency problems. First, we can examine whether managers in mutuals spend more—and therefore operate less efficiently—than managers in stockholder-owned firms. The increased spending may be due to excessive consumption of perquisites by managers, less efficient cost control, or other expense-preferring behavior. Note that such behavior also represents a tension between the two *de facto* owners of mutuals, managers and the government. Managers may prefer to inflate expenses, whereas the government prefers that the mutual reduce expenses and increase retained earnings since this improves the institution's safety and diminishes the liability of the deposit insurance fund. Second, we can ask whether mutual S&Ls have operated at output levels as efficient as those of stock S&Ls. In other words, do mutuals exploit scale economies as efficiently as stock S&Ls?

The empirical evidence sheds light on these questions. Many studies have shown that managers in mutuals exhibit expense-preference behavior relative to

10. Indeed, during 1966–1982, cash distributions to depositors were legally prohibited under FHLBB interest rate ceilings. See Masulis (1987).

11. This is achieved with the signing of perpetual proxies. These proxies can be revoked. However, disclosure requirements on the part of the S&L management are limited, the maximum number of votes a depositor can control is limited, there are restrictions on outside nominations to the board, and the board can eliminate a depositor's voting rights by simply redeeming his savings account. See Masulis (1987).

12. See Fama and Jensen (1985).

13. See O'Hara (1981).

those in stocks.[14] Moreover, other studies have found that mutuals operate at inefficient output levels relative to stocks. For example, mutual S&Ls have been found to expand deposits and loans beyond profit-maximizing levels.[15] Of course, such behavior could be motivated by a managerial desire to consume additional perquisites, so that this inefficiency could be the outcome of expense preference as well.

Another output inefficiency may be found in diseconomies of scope in mutuals. The larger the number of products the firm produces, the more complicated its management structure, and the more costly it is for owners to monitor management. Thus, managers may be tempted to expand the product offerings of their firms beyond the level at which economies of scope are maximized. There is empirical evidence that suggests that diseconomies of scope are greater in mutuals than in stocks.[16]

Choice of Ownership Structure by S&Ls

Earlier studies viewed mutual S&Ls as either cooperatives, with depositors and borrowers working for a common goal, or benevolent associations organized to encourage saving and home ownership.[17] This view was based partly on the observation that the first S&Ls were mutuals that served smaller depositors, leaving the larger ones to commercial banks and other institutions.[18] These early community-based cooperatives, which gathered deposits from the community and offered mortgages to community members, had simple operations. The fair degree of homogeneity in mortgages made it relatively easy to assess the value of the S&L's assets based on historical data. This was just as well since the absence of a *secondary market* for residual claims meant that existing and prospective owners could not rely on the information generated by capital market trading (and pricing) to assess the value of the mutuals' assets. For assets whose value is difficult to determine stock ownership is superior because the information generated by trading facilitates valuation.[19] Whereas the simplicity of the operation of S&Ls made mutuality an acceptable ownership structure, the elimination of the classic conflict between creditors (who prefer less risk) and stockholders (who prefer more) made mutuality the *preferred* structure for many S&Ls.[20] Moreover, the simplicity of the operation of S&Ls meant that managerial expertise was not a critical element in the success of S&Ls. In the early years, therefore, the S&L industry was dominated by mutuals run by managers who were not the most talented or efficient.

Over time, however, operation became more complex, and mutuals began to choose managers on the basis of expertise.[21] Moreover, the advent of deposit insurance eliminated the agency-cost-of-debt advantage of mutuals over stocks. Since their deposits are insured, depositors are indifferent to an S&L's risk-taking behavior. The agency cost of debt was essentially absorbed by the Federal Savings and Loan Insurance Corporation (FSLIC).

14. Deshmikh, Greenbaum, and Thakor (1982) make this theoretical prediction. Supporting empirical evidence can be found in Edwards (1977), Hannan and Mavinga (1980), and Smirlock and Marshall (1983).

15. See Akella and Greenbaum (1988).

16. See Mester (1991) for careful empirical documentation that stock S&Ls operate with an efficient output mix, whereas mutual S&Ls operate with significant diseconomies of scope.

17. See Hester (1968) and Brigham and Pettit (1969).

18. There are also theoretical models that suggest such a role for mutuals. See Rasmusen (1988).

19. See Fama and Jensen (1983).

20. See Meyers and Smith (1986).

21. See Masulis (1987).

Along with these developments came deregulation and an increase in competition. Mutual S&L managers have found it increasingly difficult to compete with their more efficient stockholder-owned counterparts. And their inability to augment institutional net worth through additional equity issues has made the competitive disadvantage worse. Thus, the benefits of mutuality to owners have diminished significantly. Furthermore, these increased competitive pressures mean that the probability of bankruptcy—and hence the probability of unemployment for the manager—due to inefficient behavior has increased. This means that any given level of perquisite consumption on the part of mutual managers is now more costly. Given that managers were optimally selecting their perquisites prior to deregulation, the implication is that perquisites consumption in mutuals must be *lower* after deregulation, as managers weigh the benefit of perks against the elevated probability of unemployment. Thus, the benefits of mutuality to managers have diminished as well. Combined with this is the positive incentive managers have to convert to stock ownership, since they usually benefit in the initial stock sale. The reason is that managers typically receive rights to purchase the new stock, which is usually underpriced (as in other initial public offerings). When the benefits of conversion outweigh the benefits of the new optimal (and lower) level of perquisites consumption, the S&L will convert from mutual to stock.[22] This could explain the increased number of conversions that have been witnessed in recent years,[23] as the stockholder-ownership structure has become the preferred mode for both owners and managers.

The Borrower's Choice of Finance Source

We have seen that a borrower has access to a wide array of credit sources. How does he decide which source to approach? In Figure 3.2 below, we have sketched a hierarchy of financing sources that explains the borrower's choice based on his own attributes and the resulting demand for intermediation services.[24] The borrower's financing choice in this figure tracks a typical firm's "lifecycle."

When a firm is very young, it has two striking characteristics. First, the entrepreneur in charge may be unsure of his own management expertise, so that approaching a financial intermediary that can provide this expertise is beneficial. Second, the

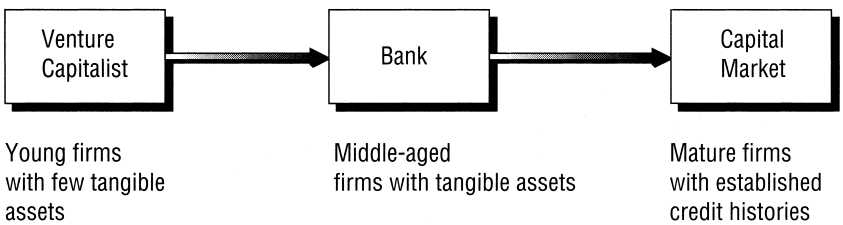


FIGURE 3.2 Hierarchy of Financing Sources

22. Mester (1991) arrives at this explanation based on her empirical analysis.

23. To convert from mutual to stock, the S&L must sell stock publicly through a standby rights offering to depositors and management, who are the eligible subscribers. The conversion plan must be first approved by two-thirds of the S&L's board of directors. If approved, it must be ratified by two-thirds of the depositors. Upon ratification, the stock can be offered to eligible subscribers, and if it is not fully subscribed, the unsubscribed portion must be sold to the public.

24. This discussion is based in part on Diamond (1989), and Chan, Siegel, and Thakor (1990). See also Boot and Thakor (1997).

borrower has few tangible assets to offer as collateral. As we will see in Chapter 5, collateral is useful in controlling moral hazard whereby borrowers either stint on effort or take excessive risks. In the absence of collateral, the lender could use equity participation as a way of addressing moral hazard. Thus, it is in the borrower's interest to seek a lender who can take an equity position and thus be able to offer capital at a "reasonable" price. Both factors suggest that such firms should go to venture capitalists.

As a firm grows and acquires tangible assets, it becomes capable of offering collateral to mitigate moral hazard. Banks, which are prohibited in the United States from taking equity positions, can now lend to such borrowers because they can offer collateral to secure their debt. Of course, all moral hazard will not be eliminated by collateral, so that there will be an important role for bank monitoring. Moreover, bank loans tend to be of short maturities, thereby generating periodic information through reassessments of the borrower. This information is reflected both in the bank's decision to renew/terminate the loan as well as in the new contract terms offered, in combination with information produced by rating agencies. This helps to reduce duplication in information production by *other* creditors of the firm, thereby diminishing overall contracting costs. The firms in this group find it better to go to banks than to venture capitalists because banks can fund their loans with insured deposits, whereas venture capitalists cannot; hence, the borrower is able to obtain a loan at a lower price.

Finally, when the firm is well-established and mature, it has a good track record for repaying its debts. This reputation can be valuable because it permits the firm to borrow at preferential rates. By taking undue asset risks, the borrower stands to lose this reputation, and thus has an incentive to limit risk-taking. Consequently, bank monitoring to combat moral hazard is less important for such borrowers, and this permits them to directly access the capital market where borrowing costs are lower; capital market access would mean that the borrower would not have to pay the bank its intermediation rents. Of course, such firms still confront problems of asymmetric information[25], so that *nondepository* financial intermediaries such as investment banks (or credit-rating agencies) play an important role in the transfer of capital from investors to such firms. This is because they make information about firms available to investors at a lower cost than they could acquire themselves. It is interesting to note that as one moves from left to right in the financing hierarchy of Figure 3.2, the intermediation services provided decline and so does the cost of credit. The venture capitalist provides financing, monitoring, and management expertise; the bank provides financing and monitoring; and the capital market provides mainly financing. Of course, this discussion is not meant to suggest that these financing sources are mutually exclusive. For example, borrowers often access the capital market for commercial paper and use banks to provide loan commitments to back up these commercial paper issues.

Blurring Distinctions Between Bank Loans and Capital Market Financing: Transaction and Relationship Loans

Although in our earlier discussion, we have characterized capital market and bank financing as distinct but sometimes overlapping choices, in recent years the distinction between these two sources of financing has become increasingly blurred. For

25. See, for example, Myers and Majluf (1984).

example, banks made syndicated loans in which multiple banks participate, and these loans are often traded in a manner similar to capital market trading. Banks make mortgage and credit card loans and then package them into portfolios, issue securities against these portfolios and sell these securities in the capital market where they are traded. This is called securitization and will be discussed in more detail in Chapter 9.

Of course, banks also make loans where they add considerable unique value and the loans are not traded. Examples are small business loans where the bank-borrower relationship has value.

Research in banking has examined the difference between loans by classifying bank loans as transaction loans and relationship loans.[26] Transaction loans include loans like credit card and mortgage loans. There is little monitoring by the bank and the loans can be repackaged and traded. The bank's value added is limited mostly to its credit analysis and standardized credit analysis before credit is extended.

Relationship loans are those where the bank generates additional value by learning about the borrower through its relationship with the borrower and providing business advice. Relationship loans offer numerous other advantages related to attenuating moral hazard and private information problems. These will be discussed in Chapter 6.

Another aspect of relationship lending that has only recently begun to be explored is that it creates the potential for *differences of opinion*. For example, a bank may judge a relationship loan to be creditworthy, but its judgment may be based on a lot of "soft," nonverifiable information. Such loans may find it difficult to obtain direct capital market financing if investors have a different (collective) opinion about the creditworthiness of the loan. In such cases, a bank—backed by sufficient capital—can act as a "beliefs bridge" between depositors/investors and borrowers and raise deposit financing to fund the relationship loan. The bank's reputation/credibility is reliably processing soft information and this may convince depositors to extend funding they otherwise may not have. This would be another contribution of banks to relationship loans.[27]

Thus, bank loans span a continuum from relationship loans at one end to transaction loans at the other. Relationship loans are the most different from capital market financing. Transaction loans are the most similar to capital market financing.

Conclusion

The process of financial intermediation is of central importance to the functioning of a modern economy. Some of the important conclusions to be drawn from our discussions are covered briefly below.

First, regulation of banks and the *raison d'etre* for the existence of banks are intertwined. Regulation is not solely the outcome of a political agenda that is separate from the reasons why banks exist. To make banking a viable business in which there is public confidence, some form of regulation is necessary. We also discussed how this regulation then becomes a component of monetary policy. Second, the incentive problems that banks and nondepository financial intermediaries resolve are such that there are natural benefits to size. Diversification can reduce incentive costs in

26. This characterization was provided by Boot and Thakor (2000). See also Rajan (1993) and Sharpe (1992) for models of relationship lending. Boot (2000) provides a review.

27. See Coval and Thakor (2005) and Song and Thakor (2006).

contracting among unequally informed agents, and information reusability is greater in larger intermediaries. Hence, financial intermediaries can derive economic benefits from being large. Third, inside (privately placed) debt has some inherent advantages over outside (publicly traded) debt because of superior access to information about the borrower that the former provides. Bank loans are inside debt. However, even within the class of contracts qualifying as inside debt, bank loans are special. The reaction of a borrowing firm's stock price to the announcement of a bank loan agreement is more favorable on average than the stock price reaction to the announcements of other forms of inside debt. Fourth, the choice of organizational form—mutual versus stock—by a depository institution depends on the interaction between a variety of factors that include differences in the efficiency with which agency problems are resolved within mutuals as opposed to stocks, the competitive environment, and the relative advantage a stockholder-owned firm has in raising capital and having complex assets priced in the capital market. This explains the initial prevalence of mutuality among thrifts and the recent trend of conversions of mutuals into stock. Finally, there is a natural hierarchy of financing sources. In its earliest phases of development, a firm has the greatest advantage in seeking venture capital, due to the (unique) ability of the venture capitalist to assist in management. At the next stage, when early survival has been accomplished, bank loans are preferred. Although banks do not assist in management to the extent that venture capitalists do, the monitoring provided by banks is of value to firms at this stage when they are still relatively small or medium-sized. Bank monitoring helps to control incentive problems within the borrowing firm. Moreover, bank loans tend to be of short maturities, thereby generating periodic information through reassessments of the borrower. This information, as well as that produced by nondepository financial intermediaries such as credit-rating agencies, helps to reduce duplication in information production by *other* creditors of the firm, and thus reduces overall contracting costs. Finally, large firms go directly to the capital market for outside debt. Bank monitoring is of lesser marginal value to such firms. However, such firms still confront problems of asymmetric information,[28] so that *nondepository* financial intermediaries such as investment banks (or credit-rating agencies) play an important role in the transfer of capital from investors to such firms. This is because they make information about firms available to investors at lower cost than they could acquire themselves.[29]

What are the implications of our analysis for *market efficiency*? Clearly, if the capital market were strong-form efficient even without financial intermediaries, the role for financial intermediaries would be extremely limited; they would at best provide some minor transactional services like "lot-breaking" of securities, that is, buying large denomination securities and selling smaller denomination claims against such securities to investors with wealth constraints. However, the theoretical and empirical results discussed in this chapter suggest two conclusions. First, given the pervasive problems of private information and moral hazard, it is reasonable to expect that credit markets are no more than semistrong form efficient, so that financial intermediaries have an important role to play in resolving information-based problems. Second, the informational efficiency of credit markets is *enhanced* by financial intermediaries, since they possess privileged financial information that is then learned by others who observe bank-borrower transactions.

28. See Ramakrishnan and Thakor (1984) and Giammarino and Lewis (1988).
29. See, for example, Diamond (1989).

Review Questions

1. Explain how a bank evolves from a primitive goldsmith and the roles played by asymmetric information and moral hazard in this evolution.
2. Can banking ever become completely deregulated? Why or why not?
3. What do we mean by a "hierarchy of financing sources"? What determines a borrower's choice of financing source?
4. Can you shed light on the following facts and explain their possible interrelationships?
   a. Commercial paper issues by nonfinancial corporations in the U.S. have grown sixfold in the last 20 years.
   b. Large money center banks are turning increasingly to "middle market" borrowers (that is, those with loan requests between \$5 million and \$200 million).
   c. Securitization has grown rapidly.
5. What is the difference between a "stock" and a "mutual"? Explain the differences in the resolutions of agency problems for these two types of organizations.
6. It has been said that the health of a nation's banking system is inversely related to the speed and efficiency of information flows in the economy. Explain.
7. In what way are banks "unique"? What is the empirical evidence on this issue?
8. What are the economic incentives for financial intermediaries to grow large?
9. How do banks help to make nonbank contracting more efficient?
10. Given below is an excerpt from "A Friendly Conversation." Comment critically on it.

    Moderator: Fine, but as long as you have fractional reserve banking, you're never going to eliminate the possibility of withdrawal risk altogether.
    Appleton: That's why you have a lender of last resort, Mike.

11. How does monetary policy affect the (short-term) growth path of an economy?
12. What are the differences between transaction and relationship loans and what is the relevance of the distinction?

Appendix 3.1 The Formal Analysis of Large Intermediaries

The Model Based on Ramakrishnan and Thakor (1984): Suppose we have assets whose owners wish to attract capital. However, there is asymmetric information about the values of these assets; the owner of each asset knows more about the value than others do. As we saw in Chapter 1, this can lead to market failure if the appropriate signals are unavailable to firms. Now suppose there are some individuals who specialize in producing information about firms at a cost. Let us imagine that there are groups of individuals, with each group specializing in producing information about a particular industry or a particular firm. The cost to an individual of producing this information is $c > 0$ and each individual is risk averse, with a utility function of $U(\bullet)$ defined over monetary wealth, that is $U(\bullet)$ is increasing and strictly concave. We assume that c is a nonmonetary cost to the information producer (i.p.); it does not figure in his utility over wealth. Moreover, it is incurred only if the i.p. actually produces information about the firm he specializes in. Also, each i.p. has a minimum level of expected utility, a $\bar{U}$ that must be guaranteed by his compensation package for producing information, or he will work in an alternative occupation.

Now suppose that the firm that wishes to attract capital (or the investor who wants to decide whether he should invest in a particular asset) approaches an i.p. directly to produce information about it and release it to the market, that is, the i.p. plays the role of a rating agency. If the i.p. is just paid a fixed fee, we have a moral hazard problem in that he will avoid actually producing information, thereby saving himself the effort-related cost c. He will simply make a quick guess, collect his fee, and send the firm on its way. Investors will recognize this and the firm's price will not move. The firm will have wasted its money.

Compensation Contracts of Individual Information Producers: But suppose the firm is able to monitor the i.p. to discover something about whether he actually invested c. This monitoring produces a signal that tells the firm about the i.p.'s effort. However, this signal is noisy. Even if the i.p. invests c in information production, the signal says that he did only with probability p. With probability $1 - p$, the signal is erroneous and indicates that the i.p. did not produce information. If the i.p. did not produce information, then the signal says that he did with probability q and that he did not with probability $1 - q$. We assume $p > q$, so that the signal is informative. Now let the i.p.'s compensation be as follows: pay him \$H if the signal says he produced information and \$L if it says he did not, with $H > L$.[1] If the i.p. does produce information, he gets an expected utility of

$$\text{EU(produce information)} = pU(H) + (1 - p)U(L) - c. \qquad [3.13]$$

If he does not produce information, he gets an expected utility of

$$\text{EU(does not produce information)} = qU(H) + (1 - q)U(L). \qquad [3.14]$$

If investors are to believe that the i.p. is credible, his compensation schedule should be incentive compatible (should induce the i.p. to invest c). That is,

$$pU(H) + (1 - p)U(L) - c \geq qU(H) + (1 - q)U(L). \qquad [3.15]$$

It also will be necessary to make sure that the i.p. is willing to work for the firm. This requires that

$$pU(H) + (1 - p)U(L) - c \geq \overline{U}. \qquad [3.16]$$

We can solve (3.15) and (3.16) to come up with H and L. We can show that in equilibrium (3.15) and (3.16) should hold as equalities, that is, treating them as equalities leads to a solution that minimizes the expected cost for each firm. To illustrate, suppose $U(x) = \sqrt{x}$ for any number x, $\bar{U} = 20$ (for simplicity), $p = 0.8$, $q = 0.2$ and $c = 10$. Solving (3.15) and (3.16) as a pair of simultaneous equations with these numbers, we get $H = 10{,}000/9$ and $L = 10{,}000/36$. The i.p. earns an expected

1. If such a compensation scheme is successful in inducing the i.p. to produce information, then it is not time consistent because everybody knows he has produced information and it is pointless to pay him less when an error-prone signal says he did not. We'll ignore this problem here.

utility of exactly 20. The expected cost of information production for each firm is $0.8\ H + 0.2\ L = 944.44$ approximately.

The Solution With an Intermediary: Now suppose that there are two i.p.s, each like the i.p. in the preceding analysis, who coalesce and form a financial intermediary of two i.p.s. Each still deals with a separate firm. However, they now pool their payoffs to avail of diversification benefits. We assume that because the i.p.s are cooperating, they can costlessly observe each other's actions. This means neither i.p. has to be concerned about his partner free-riding off his effort. So now each i.p.'s compensation becomes

$$\begin{aligned} &2H/2 = H \quad \text{if both signals are favorable} \\ &(H + L)/2 \quad \text{if only one signal is favorable} \\ &2L/2 = L \quad \text{if both signals are unfavorable} \end{aligned}$$

Assuming that signals across firms are uncorrelated, the probabilities of different compensations for each i.p. are given in the following table.

TABLE 3.5 Probabilities of Compensations

Probability of Compensation	Compensation of Each i.p.
p^2 if both i.p.s produce information and q^2 if both do not	H
$2p(1 - p)$ if both i.p.s produce information and $2q(1 - q)$ if both do not	$(H + L)/2$
$(1 - p)^2$ if both i.p.s produce information and $(1 - q)^2$ if both do not	L

Note that both i.p.s will act in concert. The firms that give them compensation contracts realize that the rules of the game have changed. They must now solve the following pair of simultaneous equations.

$$\begin{aligned} &p^2U(H) + 2p(1-p)U\left(\frac{H+L}{2}\right) + (1-P)^2U(L) - c \\ &= q^2U(H) + 2q(1-q)U\left(\frac{H+L}{2}\right) + (1-q)^2U(L) \end{aligned} \qquad [3.17]$$

and

$$p^2U(H) + 2p(1-p)U\left(\frac{H+L}{2}\right) + (1-p)^2U(L) - c = \overline{U} \qquad [3.18]$$

Generally, the solution to this will be different from the previous solution. Suppose, however, that firms continue to use the old contracts where $H = 10.000/9$ and $L = 10.000/36$. It can be checked in this case that (3.17) is satisfied exactly and that the left-hand side of (3.18) is about 20.43. That is, each i.p. in the financial intermediary enjoys a higher expected utility than he did before. Note that the expected cost of having information produced for each firm will be exactly the same as before. Thus, the formation of a financial intermediary makes i.p.s better off if firms do not alter their contracts. Of course, firms may wish to write different contracts to remove the excess utility enjoyed by the i.p.s. In this case, expected information production costs of firms are lowered.

The reason why the formation of an intermediary helps is diversification. By pooling their payoffs, the i.p.s are able to reduce individual risks. This means that they can increase their expected utility and if at least some of the benefit of this increased utility is shared with the firms they are screening, the cost of information production will also decline.

The Desirability of a Very Large Intermediary: This argument can be taken to the limit. Suppose the financial intermediary becomes infinitely large. Then, by the law of large numbers (roughly speaking) the probabilities become actual fractions. That is, if all i.p.s produce information, the intermediary knows that exactly 80 percent of them will get H each and 20 percent will get Leach. Thus, the intermediary knows that its payoff will be

$$\begin{aligned} &0.8\mathrm{H} + 0.2\mathrm{L} \\ &= 0.8\left(\frac{40000}{36}\right) + 0.2\left(\frac{10000}{36}\right) = 944.44 \end{aligned}$$

per i.p. with probability one. Since the financial intermediary itself can monitor its own members, it does not have to worry about moral hazard. Thus, it can promise each of its member i.p.s a *fixed* payment of 944.44, knowing that even though on any given i.p., it could receive either more or less than this amount, the random fluctuations around 944.44 will cancel out for the intermediary as a whole. Thus, each individual i.p.'s expected utility in this intermediary is $\mathrm{U}(944.44) - 10 = 20.73$, which is higher than with the two-i.p. intermediary passes along this gain to the firms it screens, then *information production costs are lowest with a very large intermediary*.

That is, we have shown that a *diversified information broker* can lower the cost of information production and hence the cost of exchanging capital. Once again, the pivotal function served by a financial intermediary is that of providing a more efficient resolution of informational problems.

Diversification in this model is achieved by letting each i.p. within the intermediary share the risk in the compensation of every other member i.p. That is, as we add to the size of the group, each individual compensation risk is shared by an increasing number of i.p.s. Due to the risk aversion of the member i.p.s, such diversification helps to improve welfare.[2] We shall call this "diversification by sharing risks." Another type of diversification is "diversification by adding risks."[3] In this case, a single i.p. bears 100 percent of N independent risks, with diversification occurring as N increases. This is quite different from the first form of diversification because the total wealth of the i.p. is growing as he adds more risks. That is, instead of spreading a given amount of wealth over a larger number of independent gambles, we are spreading an increasing amount of wealth over a larger number of independent gambles. Noble laureate Paul Samuelson (1963) has called such diversification "the fallacy of large numbers," because it is not generally true that, for all risk-

2. An important assumption in our analysis is that the i.p.s within the intermediary can monitor each other costlessly. Millon and Thakor (1985) show that if such monitoring is impossible, then by letting i.p.s coalesce and engage in payoff-pooling, we raise information production costs. They also show, however, that if the values of firms depend on a common, systematic element, as well as on idiosyncratic factors, then information sharing within the intermediary can lead to on overall lowering of information production costs.

3. This is considered by Diamond (1984).

averse utility functions, the individual's risk aversion toward the N^{th} independent gamble is a decreasing function of N. In other words, while a risk-averse individual would wish to take advantage of the low number of large numbers to spread a *fixed* amount of wealth over an increasingly large number of independent gambles, he would not necessarily wish to achieve such diversification at the expense of exposing an increasing amount of his wealth to the gambles. However, there are sufficient conditions involving restrictions on utility functions that such diversification is beneficial.

References

Akella, S. Rao, and Stuart I. Greenbaum, "Savings and Loan Ownership Structure and Expense-Preference," *Journal of Banking and Finance* 12, 1988, 419–437.

Allen, Franklin, "The Market for Information and the Origin of Financial Intermediation," *Journal of Financial Intermediation* 1–1, May 1990, 3–30.

Black, Fisher, "Bank Fund Management in an Efficient Market," *Journal of Financial Economics* 2, 1975, 323–339.

Boot, Arnoud, "Relationship Lending: What Do We Know?" *Journal of Financial Intermediation*, 2000.

Boot, Arnoud, and Anjan V. Thakor, "Can Relationship Banking Survive Competition?" *Journal of Finance* 55–2 April 2000, 679–714.

Boot, Arnoud, and Anjan V. Thakor, "Financial System Architecture," *Review of Financial Studies*, Fall 1997, 693–733.

Boyd, John, and Edward Prescott, "Financial Intermediary Coalitions," *Journal of Economic Theory* 38, April 1986, 211–232.

Brigham, Eugene F., and Richard R. Pettit, "Effects of Structure on Performance in the Savings and Loan Industry," in *Study of the Savings and Loan Industry* (Irwin Friend, Ed.), Federal Home Loan Bank Board, Washington, DC, 1969.

Campbell, Tim, and William Kracaw, "Information Production, Market Signalling, and the Theory of Financial Intermediation," *Journal of Finance*, September 1980, 35–4, 863–882.

Chan, Yuk-Shee, Daniel Siegel, and Anjan Thakor, "Learning, Corporate Control and Performance Requirement in Venture Capital Contracts," *International Economic Review* 31–2, May 1990, 365–381.

Coval, Josh, and Anjan V. Thakor, "Financial Intermediation as a Beliefs-Bridge Between Optimists and Pessimists," *Journal of Financial Economics* 75–3, March 2005, 535–570.

Deshmukh, Sudakar, Stuart Greenbaum, and Anjan Thakor, "Capital Accumulation and Deposit Pricing in Mutual Financial Institutions," *Journal of Financial and Quantitative Analysis* 17, December 1982, 705–725.

———, "Reputation Acquisition in Debt Markets," *Journal of Political Economy*, 97–4, 1989, 828–862.

Diamond, Douglas, "Financial Intermediation and Delegated Monitoring," *Review of Economics Studies* LI, July 1984, 393–414.

Edwards, Franklin R., "Managerial Objectives in Regulated Industries: Expense Preference Behavior in Banking," *Journal of Political Economy* 85, 1977, 147–162.

———, "Contract Costs and Financing Decisions," *Journal of Business*, 63–1–2, January 1990, S71–S91.

———, and Michael C. Jensen, "Agency Problems and Residual Claims," *Journal of Law and Economics*, 26, June 1983, 327–349.

———, and Michael C. Jensen, "Organization Forms and Investment Decisions," *Journal of Financial Economics* 14, 1985, 101–119.

Fama, Eugene, "What's Different about Banks?" *Journal of Monetary Economics* 10, 1980, 10–19.

Giammarino, Ronald, and Tracy Lewis, "A Theory of Negotiated Equity Financing," *Review of Financial Studies* 1–3, Fall 1988, 265–288.

Hannan, Timothy H., and F. Mavinga, "Expense Preference and Managerial Control: The Case of the Banking Firm," *Bell Journal of Economics* 17, 1980, 671–682.

Hester, Donald D., "Stock and Mutual Associations in the Savings and Loan Industry: A Study of the Economic Implications of Conversions," Federal Home Loan Bank Board, Washington, DC, 1968.

James, Christopher, "Some Evidence on the Uniqueness of Bank Loans," *Journal of Financial Economics* 19, 1987, 217–235.

Leland, Hayne, and David Pyle, "Informational Asymmetries, Financial Structure, and Financial Intermediation," *Journal of Finance* 32–2, May 1977, 371–387.

Lummer, Scott L., and John J. McConnell, "Further Evidence on the Bank Lending Process and the Capital-Market Response to Bank Loan Agreements," *Journal of Financial Economics* 25–1, November 1989, 99–122.

Masulis, Ronald W., "Changes in Ownership Structure Conversions of Mutual Savings and Loans to Stock Charter," *Journal of Financial Economics* 18, 1987, 29–59.

Mayers, David, and Clifford W. Smith Jr., "Ownership Structure and Control: The Mutualization of Stock Life Insurance Companies," *Journal of Financial Economics* 16, 1986, 73–98.

Mester, Loretta J., "Agency Costs Among Savings and Loans," *Journal of Financial Intermediation* 1–3, June 1991, 257–278.

———, and Anjan Thakor, "Moral Hazard and Information Sharing: A Model of Financial Information Gathering Agencies," *Journal of Finance*, 40–5, December 1985, 1403–1422.

Millon, Marcia, "*Essays on the Theory of Institutional Structures*," Unpublished PhD Dissertation, Indiana University, December 1983.

Myers, Stewart, and Nicholas Majluf, "Corporate Financing and Investment Decisions When Firms Have Information That Investors Do Not Have," *Journal of Financial Economics* 13, 1984, 187–221.

O'Hara, Maureen, "Property Rights and the Financial Firm," *Journal of Law and Economics* 23, 1981, 317–332.

Rajan, Raghuram, "Insiders and Outsiders: The Choice Between Informed and Arm's Length Debt," *Journal of Finance* 47–2, September, 1992, 1367–1400.

Ramakrishnan, Ram, and Anjan Thakor, "Information Reliability and a Theory of Financial Intermediation," *Review of Economic Studies* LI, 1984, 415–432.

Rasmusen, Eric, "Mutual Banks and Stock Banks," *Journal of Law and Economics* 31, 1988, 395–421.

Sharpe, Steven, "Asymmetic Information, Bank Lending, and Implicit Contracts: A Stylized Model of Customer Relationships," *Journal of Finance* 45–2, September 1990, 1069–1087.

PART ♦ III

Major "On-Balance-Sheet" Risks in Banking

This page intentionally left blank

CHAPTER • 4

Major Risks Faced by Banks

"Bets on the directions of interest rates are like the little girl from the nursery rhyme with the curl on her forehead. When they are good, they can be very, very good, but when they are bad, as NCNB Corp. is now finding out, they can be horrid."

Kelley Holland: *American Banker*, March 20, 1990

Glossary of Terms

OTS: Office of Thrift Supervision. This is a national regulatory agency for the thrift industry.

Zero-coupon bonds: Bonds that pay no coupon, so that the entire repayment to bondholders is at maturity.

Immunization: The act of insulating the institution from interest rate risk.

Going Long: Purchasing a security.

Going Short: Selling a security without owning it.

Introduction

Risk is endemic to business but central to banking. What precisely do we mean by risk? In the context of business, risk is the distillate of randomness in the process by which earnings are generated. This randomness may be avoidable, in large part, in which case the risk is voluntarily accepted, perhaps even sought, as routine business decision; hence a "businessman's risk." Alternatively, the risk may be unavoidable, as in the case of a *force majeure* or an "act of god," in which case the only protection is to seek outside insurance or to exit the industry. The risks in business are as diverse as life itself. The businessman faces possible losses owing to flood, plague, fire, machine failure, worker alienation, sabotage, war, or capricious acts of government that destroy or appropriate property (sovereign risk). Shoe stores as well as financial intermediaries face all of these risks, but risks of the avoidable variety define the business of banking.

It is important to bear in mind that risk is *not* due to *variability per se*, but rather due to *uncertainty*. In an *ex post* sense, we often use the terms variability and uncertainty synonymously. However, in an *ex ante* sense, the two are quite distinct. We can have a cash flow, for example, that is known *for sure ex ante* to be 1, –100, 1,000, and 0 in years 1, 2, 3, and 4, respectively. This cash flow has a very high intertemporal variability, but has *no* risk. By contrast, a cash flow that can be either $+1$ or -1 with equal probability in each of the next 4 years has less intertemporal variability but more risk.[1] Risk, then, is related to uncertainty or lack of predictability.

The Source of Business Risk

What kinds of risks do banks face? To address this question, it is important to note that banks are essentially no different from other firms when it comes to the *raison d'etre* for being exposed to risk. A bank's shareholders, or the shareholders of any other firm for that matter, bear risk when the economic nature of the firm's "assets" is somehow different from that of its "liabilities."

Consider a steel fabrication company in Figure 4.1 below.

In Figure 4.1, the risk to the fabricator's shareholders arises primarily from the fact that the prices of raw steel and fabricated steel do not move in perfect unison. This exposes the fabricator's profit margin to random fluctuations and creates risk for its shareholders. Note that this risk comes from a mismatch on the fabricator's "balance sheet." Its liability (what it owes its suppliers for raw steel) is of a different nature from its "assets" (the fabricated steel it sells to its customers) because the prices of raw and fabricated steel are not perfectly correlated.

Now suppose the fabricator purchases its raw steel in Japan, paying its suppliers in Japanese yen, and sells fabricated steel in the United States, receiving dollars from its customers. In this case, we see that the fabricator's balance sheet is even more mismatched because of the different currencies involved. Consequently, its

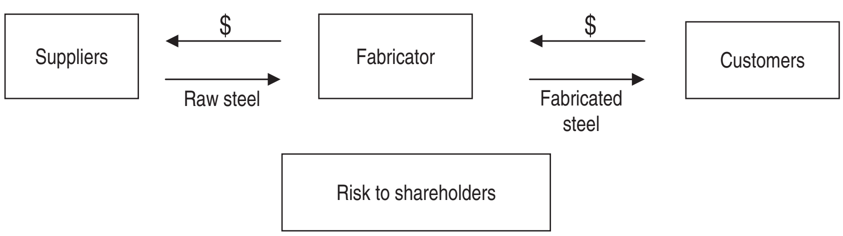


FIGURE 4.1 Risks Faced by a Domestic Steel Fabricator

1. For the havoc caused by not distinguishing between variability and risk, see Sprenkle and Miller (1980).

shareholders are exposed to even more risk. In particular, they face currency risk (due to the lack of perfect correlation between movements in the yen and the dollar) in addition to the price risk they faced earlier.

In general then, *mismatches* imply risks. This is a notion familiar to us from Chapter 2. Qualitative asset transformation involves mismatching the two sides of the balance sheet and, hence, creates risk. What are the major mismatches for banks? These are described in Figure 4.2.

In Figure 4.2 we see that a typical bank's assets (e.g., loans) and liabilities (e.g., demand deposits) are mismatched along three dimensions. First, the assets usually involve greater credit risk than the liabilities, i.e., the bank's claim against the borrower is riskier than the depositor's claim against the bank. Second, the assets are usually of longer maturity than the liabilities. For example, a loan may have a 1-year maturity, whereas demand deposits are withdrawable on demand (zero maturity). This creates interest rate risk. Third, a bank's liabilities are usually more liquid than its assets, i.e., a depositor is able to withdraw his deposits without notice, whereas the bank cannot call back a performing loan at-will and the loan may also not trade in an active market. This creates liquidity risk. We shall now discuss each of these risks in more detail.

Credit, Interest Rate, and Liquidity Risks

1. **Default or Credit Risk:** This is the risk that a party with whom you contract fails to fully discharge the terms of the contract. For a bank, this is the risk that a borrower fails to make the contractual payment on a timely basis. This kind of risk is central to virtually all rental transactions, and as in the case of almost all insurance contracts, moral hazard is a key element in default risk.

   The avoidability of default risk has two aspects. Banks can choose assets with little or no default risk, such as government securities or the debt of triple-A rated borrowers. Such a strategy, however, may provide a return only slightly, if at all, greater than the bank's cost of borrowing, and such a low (albeit relatively safe) profit margin may be unattractive to the bank.

   Given that the bank chooses assets with substantial default risk, its ability to control default risk derives from its ability to resolve moral hazard and other

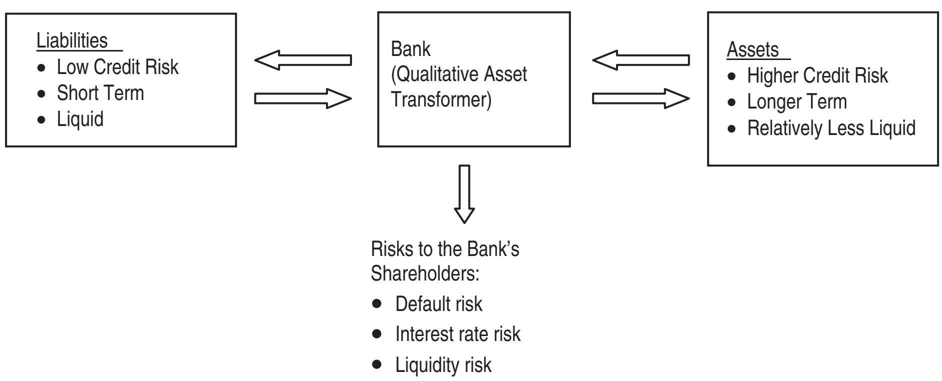


FIGURE 4.2 Major Mismatches for Banks

informational problems. In our earlier discussion in Chapters 2 and 3, we argued that banks enjoy a special advantage in screening and monitoring borrowers. However, it is virtually impossible to monitor a borrower so closely that default (or credit) risk can be completely eliminated.

There are two sources of default risk: cash flow variations beyond the borrower's control (physical hazard) and moral hazard. As for the first source of default risk, the bank's role as a financial intermediary is to *screen* the borrower so that it can accurately assess the risk it is taking in lending. This involves an analysis of the borrower's financial statements and other relevant financial and operating information about the borrower. In this capacity, the bank only assesses the risk but does not bear it, that is, it is acting as a pure broker. Thus, its role is similar to that of a bond rating agency or an investment banker. Our discussion in Chapter 3 suggests that large (diversified) information brokers can motivate their members to produce information at lower cost than is possible without intermediation. Thus, the bank should be able to efficiently generate information about default risk stemming from cash flow variations beyond the borrower's control.

As for moral hazard, the bank's monitoring capability is important. As we will see in some detail in the next chapter, the borrower has an incentive to take actions after taking a (risky) loan that increase the bank's risk exposure. This is why covenants are included in loan contracts to restrict the activities of borrowers. However, bank monitoring of borrower compliance with these covenants is important to control moral hazard. Thus, the efficiency with which the bank performs its basic functions as an FI is a key determinant of its own credit risk exposure. Moreover, loans are subject to management as a portfolio. A bank can control its default risk by holding in its asset portfolio many loans with imperfectly correlated prospects and thereby diversifying across loans.

2. **Interest Rate Risk:** This risk derives from variation of market prices. If the firm's assets and liabilities are traded, they are subject to being revalued by the market. Any such revaluation, due to changes in either the level or structure of interest rates, is described as interest rate risk. Let's consider a simple example. Suppose a bank makes a 2-year, $1 million loan for which it charges 10 percent interest. It faces the choice of financing the loan with a 2-year deposit at 9 percent per annum, or with a 1-year deposit at 8 percent per annum. The former choice will result in $10,000 in *certain* interest earnings for each of the 2 years. However, if the bank chooses the 1-year financing, it will earn $20,000 in year 1, but its earnings in year 2 will depend on the currently unknown 1-year interest rate that will prevail a year from now. Should the 1-year rate remain unchanged, the bank will enjoy a second year of earning $20,000. And if the 1-year rate were to fall to 5 percent, management will do even better and record second-year earnings of $50,000. But interest rates rise too, as the S&L industry discovered to its chagrin in 1980–81, and should the 1-year rate rise to, say, 12 percent, the bank will sustain a loss of $20,000 in year 2. This example illustrates both the substance of interest rate risk and its discretionary aspect. The risk could have been avoided with the choice of 2-year financing, assuming, of course, that 2-year financing was available. If not available, or if available only at a rate exceeding 10 percent, the bank need not have offered the borrower a 2-year fixed-rate loan.

Another aspect of interest rate risk, from the standpoint of the bank, is *prepayment risk*. This risk arises from the borrower's *option* to prepay. If interest rates rise, there will be no prepayment. But if interest rates fall sufficiently after the loan has been taken, the borrower is likely to prepay the loan by taking advantage of refinancing at a lower rate.

3. Liquidity (Withdrawal) Risk: This is the risk that an asset owner (seller of a house or a borrower selling its indebtedness) will not be able to realize the full value of that asset at the time a sale is desired. In banking, the liquidity risk faced by a borrower is that the lender may choose *not* to renew a loan that a borrower wants to renew. Similarly, the liquidity risk faced by a bank is that depositors may unexpectedly withdraw their deposits and the bank may be unable to replace them without impairing its net worth. This risk applies symmetrically to borrowers in their relationship to banks, and to banks in their relationship to depositors. The most extreme manifestation of liquidity risk is that the seller of the asset is simply *unable* to sell the asset at any price. In credit markets, this phenomenon is known as *credit rationing*, whereby a borrower is refused credit regardless of the price it is willing to pay. We shall have more to say about credit rationing in Chapter 6, but suffice to say this phenomenon has long perplexed economists in particular because it indicates an apparent suspension of price as the arbiter of allocations.

Liquidity risk has yet another interpretation. Asset markets vary widely in their development and level of activity. At one extreme, we have flea markets for "one-of-a-kind" antiques of dubious authenticity. At the other we have 24-hour around-the-world markets for currencies and government debt in which large quantities are traded at relatively low cost. More primitive and less active markets are typically characterized by large bid-ask spreads, where the *bid-ask spread* is defined as the difference between the price at which one can buy a security and the price at which one can sell it at the same place and time. For example, you can *buy* a Treasury bill at an *ask* of $\$98\frac{1}{2}$ and *sell* it at a *bid* of $\$98\frac{1}{4}$, in which case the bid-ask spread is $\$98\frac{1}{2} - \$98\frac{1}{4} = \$\frac{1}{4}$. Bid-ask spreads range from small fractions of a percent of the asset value for actively traded assets, to 6 or 7 percent for residential property. Still larger bid-ask spreads hold for infrequently traded, heterogeneous, and hard-to-value objects. Bid-ask spreads are the cost of simultaneous purchase and sale of an asset, and reflect the liquidity in asset markets.

Illiquid assets are those for which "full value" is not readily realizable. That is, time and effort are required to realize the full value of an asset that is relatively illiquid.[2] Hence, a bank holding illiquid assets can find itself unable to redeem its liabilities on short notice, and the problem of managing the balance sheet against this eventuality is referred to as liquidity or cash management (cash is the asset with liquidity *par excellence*). The central bank, with its capacious lender-of-last-resort facility, was created to address those instances when the bank, having sound albeit illiquid assets, is unable to meet its withdrawals. The central bank provides the bank with crisis-avoiding liquidity by lending to the bank against its illiquid but otherwise presumably sound earning assets. Indeed, the central bank was designed to socialize a portion of the bank's liquidity problem.

In the remainder of this chapter we shall address interest rate and liquidity risks in greater detail. Default risk will be considered in Chapters 5 and 6, where lending will be the focus. The rest of this chapter is organized as follows. First,

2. Were all assets perfectly liquid, there would be no role for marketing.

we analyze the *term structure* of *interest rates* and discuss how the term structure is determined under certainty and uncertainty. We then discuss the concepts of *duration* and *convexity*. These concepts are basic to the notion of interest rate risk, so it is important to understand them before we discuss interest rate risk in detail, which we do next. Selected interest rate risk management techniques are subsequently examined. Next, we turn to liquidity risk, followed by concluding remarks. A case study is provided to illustrate some practical issues in interest rate risk management.

The Term Structure of Interest Rates

Review of Fixed-Income Valuation

What is the current value of a \$250 riskless cash flow to be received in 1 year? We solve this problem by using the principle of *riskless arbitrage*. In particular, to prevent riskless arbitrage – which is essential in an efficient capital market – the price of this riskless cash flow in equilibrium must be related to the prices of other riskless instruments. In particular, suppose we observe that a United States government bond that promises \$100 in 1 year is currently trading at \$94.56. From this, we can deduce that the implicit 1-year return on riskless instrument is 5.75 percent (since $\$94.56\,[1 + 0.0575] = \$100$). Thus, we should be currently willing to pay $\$250/[1.0575] = \$236.41$ for the riskless promise to receive \$250 in 1 year.

But what if the riskless cash flow is promised to us 2 years from now? Well, then we have to find a riskless instrument of similar maturity (2 years) and payment characteristics (the only promised payment is 2 years from now and there are no interim payments). Suppose we observe that United States government "pure-discount" bonds with a 2-year maturity that promises a \$100 payment are currently trading at \$88.58. Then we can deduce that the 2-year riskless yield, on an annualized basis, is given by, i_o^2, where $\$100/\left[1 + i_o^2\right]^2 = \$88.58$. Solving this equation implies an annual two-period yield of $i_o^2 = 6.25$ percent. Thus, we get Figure 4.3.

That is, even though both the year 1 and year 2 cash flows are riskless, they have different discount rates applied to them. Why?

The reason is that future one-period interest rates are expected to *increase*. In our example, we know that the 1-year riskless rate at date 0 is 5.75 percent and the 2-year riskless rate at date 0 is 6.25 percent. We can infer the 1-year riskless interest rate, i_1^1, that is expected to prevail in the future at date 1. We can solve for it as follows:

$$\$221.45 = \frac{\$250}{[1.0575]\,[1 + i_1^1]}$$

which yields $i_1^1 = 6.75$ percent. That is, the two-period rate 6.25 percent is the *geometric average* of the successive one-period rates, 5.75 percent and 6.75 percent.

	0	1	2
Cash Flow		$250	$250
Discount Rate		5.75%	6.25%
Present Value at t = 0		$236.41	$221.45

FIGURE 4.3 Cash Flows and Discount Rates

The Yield Curve

What we have seen above is that interest rates on debt instruments of different maturities are related through investors' expectations about future interest rates. Our discussion deals with zero-coupon (pure-discount) bonds until we get to duration. A useful concept for this discussion is *yield to maturity* (YTM), which is defined as the internal rate of return that equates the present value of the future cash flows from a bond to the current market price of the bond. The relationships among the yields on different bonds are summarized by the *term structure of interest rates*. We define the term structure of interest rates (or the yield curve) as the relationship between the YTM and the length of time to maturity for *debt instruments of identical default risk characteristics*. It is critical to equalize the default risk of the bonds whose yields we are comparing. For simplicity, we will confine our attention to bonds without default risk. Thus, the *YTM* on a bond with m periods to maturity is defined as the annualized equivalent discount rate at which the cash flows from the bond must be discounted m periods to arrive at its market price. Figures 4.4 and 4.5 show two different yield curves, each describing the yields of bonds that are identical, except in maturity. The yield curve in Figure 4.4 is for U.S. Treasuries and is upward sloping. It is the "on the run" curve, in which the *implicit* zero-coupon yield curve is interpolated from full-coupon bond prices. The yield curve in Figure 4.5 is for German government securities. It is cup shaped. For shorter maturities, this yield curve is "inverted," that is, the YTM decreases with maturity. For intermediate maturities, it is virtually flat, that is, the YTM is almost independent of maturity in this range. And for longer maturities, the yield curve slopes upward, that is, the YTM rises with maturity.

What determines the shape of the yield curve? For simplicity, we will examine this question first in a world of perfect certainty. Uncertainty will be dealt with subsequently. In both cases we assume that a financial market equilibrium precludes *riskless arbitrage*.

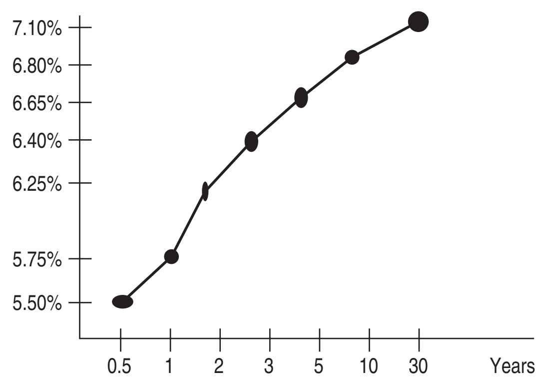


FIGURE 4.4 Risk-Free Term Structure for U.S. Treasury Securities as of July 25, 1996

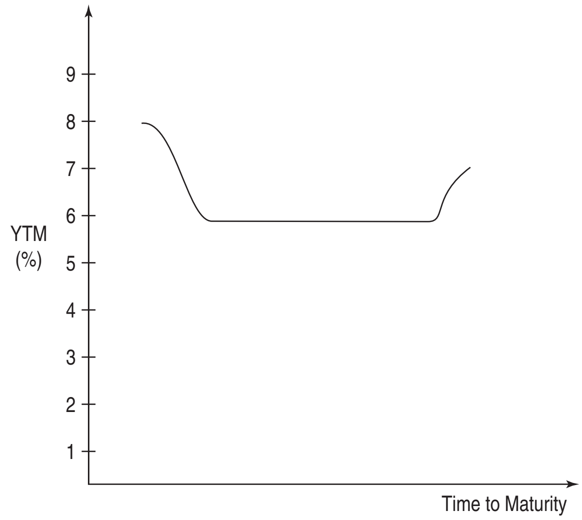


FIGURE 4.5 Yield Curve for Government Securities in Germany as of March 22, 1993

Yield Curve Determination Under Certainty

The Basic Model: Let P_t^m and i_t^m be the price and YTM, respectively, at time t of a bond of maturity m years. We assume the unit of time is 1 year, and all bonds are traded, so that prices are available from the market. As an illustration, we will examine the yield relationship between two bonds, one with a maturity of 1 year and the other with a maturity of 2 years. For simplicity, we will assume that each is a zero-coupon (pure-discount) bond and has a face value, F, of \$1. A zero-coupon bond makes a single promised payment (often called a balloon payment) at maturity, and no payments prior to that. Now, the YTM on the 1-year bond at the present time ($t = 0$), i_0^1, is the internal rate of return that discounts the \$1 face value over one period to equal the current market price of the bond.

$$P_0^1 = \frac{F}{1 + YTM} = \frac{1}{1 + i_0^1}. \tag{4.1}$$

Similarly, the YTM on the 2-year bond at $t = 0$, i_0^2, is the internal rate of return that discounts the \$1 face value over two periods to equal the current market price of the bond.

$$P_0^2 = \frac{F}{(1 + YTM)^2} = \frac{1}{\left(1 + i_0^2\right)^2}. \tag{4.2}$$

Now suppose we take \$1 today and invest it in the 2-year bond. Because it sells at $\$P_0^2$, we will be able to buy $1/P_0^2$ units of it. Then, 2 years from now (at $t = 2$), our investment will fetch us a (sure) payoff equal to the number of bonds we have bought $\left(1/P_0^2\right)$ times the face value of each bond (\$1). That is, our payoff at $t = 2$ will be [using (4.2)]

$$1/P_0^2 = \left(1 + i_0^2\right)^2. \tag{4.3}$$

Another use of our \$1 would be to invest it in the 1 year bond right now. We will be able to buy $1/P_0^1$ units of it at $t = 1$, then our payoff will be the number of bonds we have bought $\left(1/p_0^1\right)$ times the face value of each bond (\$1). That is, our payoff at $t = 1$ will be [using (4.1)]

$$1/P_0^2 = \left(1 + i_0^1\right). \tag{4.4}$$

What shall we do with this money at $t = 1$? Invest it, of course! Suppose we invest in another zero-coupon, \$1 face value, 1-year bond that will be issued a year from now (or equivalently, a multiyear bond with 1 year left to mature). Since we are currently in a world of certainty, we should be able to forecast the price, P_1^1, of this 1-year bond (issued 1 year from now) with perfect accuracy. With $\$\left(1 + i_0^1\right)$ to invest, we should be able to buy $\left(1 + i_0^1\right)/P_1^1$ units of this bond. Note that the YTM, i_1^1, of this bond is the internal rate of return that discounts the \$1 face value over one period to equal the current bond market price, and is thus

$$P_1^1 = 1/\left(1 + i_1^1\right). \tag{4.5}$$

Since we have bought $\left(1 + i_0^1\right)/P_1^1$ units of this bond at $t = 1$, and the face value of each unit is \$1, our payoff at $t = 2$ will be [using (4.5)]

$$\left[\left(1 + i_0^1\right)/P_1^1\right] \times 1 = \left(1 + i_0^1\right)\left(1 + i_1^1\right). \tag{4.6}$$

The Absence of Arbitrage and the Yield to Maturity Relationship: Equilibrium in this market requires that there be no riskless arbitrage opportunities. That is, we should not be able to do better at $t = 0$ with either the strategy of investing into the 2-year bond or investing in the 1-year bond and rolling over the proceeds into another 1-year bond. Both strategies should yield identical proceeds at $t = 2$ since we started out in each with identical \$1 investments. That is, the expressions in (4.3) and (4.6) should be equal. This gives

$$\left(1 + i_0^2\right)^2 = \left(1 + i_0^1\right)\left(1 + i_1^1\right),$$

or

$$\left(1 + i_0^2\right) = \sqrt{\left(1 + i_0^1\right)\left(1 + i_1^1\right)}. \tag{4.7}$$

Thus, the (annualized) YTM on the 2-year bond should be the *geometric average* of the YTMs on two successive bonds, each of maturity 1 year. This relationship is sometimes known as the *expectations hypothesis*, because it says that the yield on a long-term bond should be based on the expectations of investors about the yields on a sequence of short-term bonds. The general form of (4.7) for any arbitrary number of years, n, is

$$\left(1 + i_0^n\right) = \sqrt[n]{\left(1 + i_0^1\right)\left(1 + i_1^1\right)\left(1 + i_2^1\right)\left(1 + i_3^1\right)\ldots\left(1 + i_{n-1}^1\right)} \tag{4.8}$$

Spot Rates and Forward Rates: The future yields, i_1^1, i_2^1, i_3^1, are known as *forward* rates, whereas the current yields, $i_0^1, i_0^2, \ldots, i_0^n$, are known as *spot* rates. Note that the forward rate for any period in the future can be defined with the help of a ratio of bond prices. To see this, solve (4.7) to obtain

$$i_1^1 = \frac{\left(1 + i_0^2\right)^2}{\left(1 + i_0^1\right)} - 1.$$

Now, substituting for $1 + i_0^1$ and $1 + i_0^2$ from (4.1) and (4.2) respectively, we get

$$i_1^1 = \frac{P_0^1}{P_0^2} - 1.$$

Similarly, we can obtain $i_2^1 = \frac{P_0^2}{P_0^3} - 1$, and so on. A one-period-hence forward rate can thus be thought of as the interest rate on a one-period loan starting at some future point in time. An n-period-hence forward rate is the interest rate on an n-period loan starting at some future point in time. The general formula for the YTM on a bond of maturity n periods to be issued t periods from now (that is, the n-periods hence forward rate for time t) is $i_t^n = \sqrt[n]{\frac{P_0^t}{P_0^{n+t}}} - 1$. We can see now how the shape of the yield curve is determined. If investors believe that short-term interest rates will keep rising, then $i_0^1 < i_1^1 < i_2^1 < \ldots < i_{n-1}^1$, so that $i_0^1 < i_0^2 < i_0^3 < \ldots < i_0^n$, and the yield curve will be upward sloping. On the other hand, if investors believe that short-term interest rates will keep falling, then the yield curve will be inverted, or downward sloping. Given a set of bond prices, we can compute the implied forward rates in the market as we do in the example below.

Notice that the geometric mean of 5 percent, 9.03809 percent, and 16.25469 percent equals the current 3-year yield of 10 percent. Likewise, the geometric mean

Example 4.1 Suppose there are three zero-coupon bonds that are identical in all respects except maturity. Each bond has a face value of \$10 million. One of them matures a year from now and is currently selling at \$9,523,809. The other matures 2 years from now and is currently selling at \$8,734,386. The third matures 3 years from now and is currently selling at \$7,513,148. Compute the YTM for each of the three bonds, plot the yield curve (assuming that you can interpolate smoothly), and compute the available forward rates.

Solution We will solve this problem in two steps. First, we will use the specified bond prices to compute the various date-zero YTMs. Second, we will calculate the implied forward rates for different maturities by computing ratios of bond prices.

Step 1 Using our previous analysis, we have

9,523,809 = 10,000,000/(1 + i_0^1), which gives $i_0^1 = 0.05$ or 5 percent.

Similarly,

$8{,}734{,}386 = 10{,}000{,}000/(1 + i_0^2)^2$, which gives $i_0^2 = 0.07$ or 7 percent. And,
$7{,}513{,}148 = 10{,}000{,}000/(1 + i_0^3)^3$, which gives $i_0^3 = 0.10$ or 10 percent.

Step 2 We will now compute the implied forward rates. The data given to us are that $P_0^1 = \$9{,}523{,}809$, $P_0^2 = \$8{,}734{,}386$, and $P_0^3 = \$7{,}513{,}148$. Now,

$$\begin{aligned} i_1^1 &= \frac{P_0^1}{P_0^2} - 1 \\ &= \frac{9{,}523{,}809}{8{,}734{,}836} - 1 \\ &= 9.03809\%, \end{aligned}$$

and

$$\begin{aligned} i_2^1 &= \frac{P_0^2}{P_0^3} - 1 \\ &= \frac{8{,}734{,}836}{7{,}513{,}148} - 1 \\ &= 16.25469\%. \end{aligned}$$

of 5 percent and 9.03809 percent equals the current 2-year yield of 7 percent. In addition, the mean of the current 2-year yield of 7 percent and the 1-year rate 2 years hence of 16.25469 percent will equal the current 3 year-rate of 10 percent. Thus, all possible 3-year investment strategies should produce identical returns. Our analysis thus far has proceeded under the assumption of certainty. We now introduce uncertainty about future interest rates.

The Lure of Interest Rate Risk and Its Potential Impact

As we saw in our earlier examples, yields of bonds of different maturities can be different. In Figure 4.3 we depicted a case in which the 1-year yield is 5.75 percent and the 2-year yield is 6.25 percent. That is, if we buy the 1-year bond at date 0 and hold it until date 1, we get a return of 5.75 percent and if we buy the 2-year bond at date 0 and hold it until maturity at date 2, it will give us a return of 6.25 percent. The difference in returns, 6.25 percent−5.75 percent = 0.5 percent, is called the *term premium*. We may define an *m-period term premium* as the difference between the expected return on holding for a one period of a bond with maturity $m + 1$ periods at the time of purchase and the return on a bond of a one-period maturity. If term premiums are positive, then longer-term bonds should have higher expected returns. In a world of certainty, the term premium reflects simply investors' expectation that future interest rates will be higher than current rates. But in a world of uncertainty—in which interest rates fluctuate randomly—the term premium has two components:

one reflecting *expected* changes in future interest rates, and the other reflecting a premium demanded by risk-averse investors for bearing the risk (in holding longer maturity bonds) that future changes in interest rates will deviate from what is expected (this can be viewed as a premium for bearing interest rate risk).

The term premium is usually positive. This can be seen in Figure 4.6 below, which depicts the estimated 10-year term premium in the United States Treasury Bond market. This figure shows that term premiums have declined since 1990 and have fallen sharply since 2004. This suggests a greater willingness on the part of investors to hold longer maturity securities. Given investor risk aversion, this may be indicative of a lower perceived macroeconomic volatility.

Evidence of a positive term premium can also be seen in Table 4.1, which provides data on government bond yields in different countries.

The term premium is usually positive and creates a strong inducement for banks to mismatch their asset and liability maturity structures. By holding assets of longer maturities than their liabilities, banks can profit from a positive term premiums. This is the lure of interest rate risk. But this is risky too, as the following examples shows.

FIGURE 4.6 United States – Estimated Ten-Year Term Premium in the U.S. Treasury Market, 1990–2005

Note: Estimated instantaneous term premium at ten-year maturiy.

Sources: Don H. Kim and Jonathan H. Wright, "An Arbitrage-Factor Three-Factor Term Structure Model and the Recent Behavior of Long-Term Yields and Distant-Horizon Forward Rates," Federal Reserve Board, Finance and Economics Discussion Series Number 2005–33, August 2005; and the Federal Reserve.

TABLE 4.1 Government Bond Yields as of December, 2005

Country	2-year yield	10-year yield
United States	4.65%	4.8%
Euro Area	2.9%	3.5%
United Kingdom	4.42%	4.54%
Japan	0.20%	1.70%

Source: JP Morgan Economic Research, November 18, 2005.

Example 4.2 Suppose a bank's only asset is a 5 year United States government zero-coupon bond that promises to pay $100 million in 5 years. Its only liability is a 1-year $100 million certificate of deposit (CD). The yield to maturity (YTM) on 1-year riskless instruments is 5.75 percent and on 5-year riskless instruments is 6.65 percent.

This bank's balance sheet in economic value terms will look like this:

Economic Value Balance Sheet (in millions)

Assets		**Liabilities and Equity**	
Government bond	$72.48	CD	$70.92
		Equity	$1.56
Total	$72.48	Total	$72.48

The economic value of the government bond is $\$72.8 = \dfrac{\$100}{(1.10665)^5}$ whereas the economic value of the CD is $\$70.92 = \dfrac{\$100}{1.0575}$.

The economic value of the bank's equity is a plug and it arises from the term premium represented by the difference in the rates or return on the bank's assets and liabilities. As long as interest rates do not change, the bank will earn the term premium.

Now what happens to the value of the bank's equity if there is a parallel shift of the yield curve and all yields increase by 100 basis points? The new economic value balance sheet now looks like this:

Assets		**Liabilities**	
Government Bonds	$69.17	CD	$70.26
		Equity	−$1.09
Total	$69.17	Total	$69.17

The new economic value of the government bond is $\dfrac{\$100}{[1.0765]^5} = \$69.17$ and the new economic value of the CD is $\dfrac{\$100}{[1.0675]} = \$70.26$.

The equity value, which is a plug, is value of assets – value of liabilities = $\$69.17 - \$70.26 = -\$1.09$.

So we see that even though there was only a modest and equal increase in all interest rates, the economic value of equity fell from $1.56 million to a negative $1.09 million. Why? The reason is that the long-term cash flow represented by the bank's asset has a value that is much more sensitive to interest rate changes than the short-term cash flow represented by the bank's liability. Thus, banks that are typically mismatched in a manner similar to our hypothetical bank – with assets of longer maturity than liabilities – experienced a decline in their equity values when interest rates rise. This kind of interest rate risk arises because a typical bank's assets and liabilities are mismatched in a particular way.

The existence of a positive term premium has profound implications for banks. On the one hand, it allows banks to profit from a maturity mismatch on their balance sheets. On the other hand, it imposes interest rate risk on banks. So, while the lure of profiting from maturity mismatching can be quite strong, the risk of mismatching can be ruinous, as many S&Ls and Orange County, CA, found out to their chagrin.

Could the bank have hedged its shareholders against interest rate risk by matching maturities? Not necessarily. The reason is that the banks need to match the *exact timing* of their asset and liability cash flows. Shorter-term cash flows behave differently than longer-term cash flows. To hedge its shareholders against interest rate risk, the bank must understand something about how asset and liability values will change, *given* changes in market yields. That is, the bank's shareholders will be protected against interest rate movements if, *for a given change in market yields*,

Percentage Price Change in Assets = Percentage Price Change in Liabilities
or

$$\left.\frac{\Delta P_A}{P_A}\right|_{\Delta_i} = \left.\frac{\Delta P_L}{P_L}\right|_{\Delta_i} \tag{4.9}$$

where ΔP_A = change in price of asset, P_A = price of asset, ΔP_L = change in price of liability, P_L = price of liability, and Δ_i = change in interest rate.

Let us now examine the value $\left.\frac{\Delta P_A}{P}\right|\Delta_i$.

Consider first a *flat* term structure, with $i = 10$ percent and a 10-year zero-coupon bond with \$100 par. How will the price of this bond change if yields (interest rates) change by 1 basic point?

$$P(\text{no change}) = \frac{\$100}{(1.10)^{10}} = \$38.5543$$

$$P|_{\Delta_i = +0.0001} = \frac{\$100}{(1.1001)^{10}} = \$38.5193$$

$$P|_{\Delta_i = -0.0001} = \frac{\$100}{(1.10009)^{10}} = \$38.5894$$

$$\left.\frac{\Delta P}{P}\right|_{\Delta_i = +0.0001} = -0.09\%$$

$$\left.\frac{\Delta P}{P}\right|_{\Delta_i = -0.0001} = 0.09\%$$

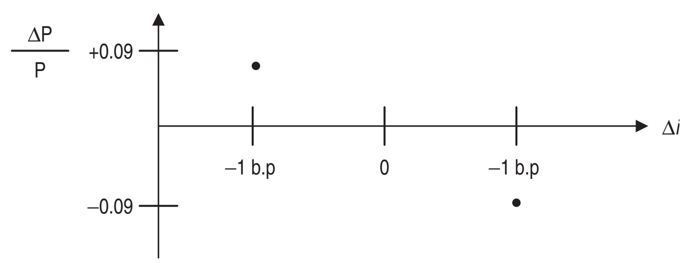


FIGURE 4.7 Price Changes for 1 Basis Point Change in Yields

Duration

The Inappropriateness of Maturity for Coupon-Paying Bonds

We saw that relative price change ($\Delta P/P$) is related to the yield change (ΔR). A mathematical relationship between $\Delta P/P$ and ΔR is given by *duration*, which is related to but different from maturity. The *maturity* of a bond tells the investor how long he must wait before receiving the terminal cash flow of the bond, or alternatively when the bond will mature or be redeemed. The maturity of a bond, however, does not give the investor all the needed information about the price volatility of the bond, unless it is a zero-coupon bond. This is because bonds of the same maturity can differ in their coupon payments through time. Moreover, in addition to coupon payments, bonds often provide other cash flows before maturity, such as amortizations. A bond that makes relatively large coupon payments early or amortizes rapidly has a shorter effective maturity than a bond that makes most of its large coupon payments late in the life of the bond. The reason is that the former generates much of its total cash flow well before its actual maturity date, whereas the latter skews its cash flows closer to its actual maturity date. We should, therefore, expect different sensitivities of the prices of these bonds to changes in interest rates. Note that we are now shifting our focus from zero-coupon bonds to bonds that may or may not pay coupons. All bonds we consider in our analysis are *nonamortizing*, that is, only coupon payments are received prior to maturity, and the entire principal is paid at maturity.

Duration Is the Answer

Duration, which is calibrated in the same temporal units as maturity, captures the timing of *all* cash flows generated by a bond, not just the terminal cash flow, and therefore is a more sophisticated measure of cash flow timing.[3] The duration of a bond is defined as the weighted average of the times to arrival of *all* scheduled future payments of a bond, where the weight attached to each payment reflects the relative contribution of that payment to the value of the bond. That is, each weighting factor is the present value of that payment divided by the present value of all payments of the bond. Consider a bond with N years to maturity, coupon payments $C_1, C_2, \ldots C_N$ where C_t is the coupon paid t years from now, and a principal (balloon) payment of B_N made at maturity. Let the term structure be *flat*, with *i* as the annual yield for all cash flows. Then the price of the bond at $t = 0$ is the present value of future payments:

$$P = \frac{C_1}{1+i} + \frac{C_2}{(1+i)^2} + \cdots + \frac{C_N + B_N}{(1+i)^N} \tag{4.10}$$

To see how P is related to R, let's take a derivative

$$\frac{dP}{di} = \frac{-C_1}{(1+i)^2} + \left[\frac{-2C_2}{(1+i)^3}\right] + \ldots + \left[\frac{-N[C_N + B_N]}{(1+i)^{N+1}}\right]$$

3. This concept was introduced by Macaulay (1938). Our treatment relies in part on generalizations by Fisher and Weil (1971), and Ingersoll, Skeleton, and Weil (1978).

or

$$dP = -\frac{-di}{1+i} + \left[\frac{C_1}{(1+i)} + \frac{2C_2}{(1+i)^2} + \ldots + \frac{N(C_N + B_N)}{(1+i)^N}\right]$$

Dividing both sides by P gives us:

$$\frac{dP}{P} = \frac{-di}{1+i}\left[\frac{\frac{C_1}{(1+i)} + \frac{2C_2}{(1+i)^2} + \ldots + \frac{N(C_N+B_N)}{(1+i)^N}}{\frac{C_1}{(1+i)} + \frac{C_2}{(1+i)^2} + \ldots + \frac{(C_N+B_N)}{(1+i)^N}}\right]$$

We can write this as:

$$\frac{dP}{P} = \frac{-di}{(1+i)}\left[1\left\{\frac{\frac{C_1}{(1+i)}}{\frac{C_1}{[1+i]} + \frac{C_2}{[1+i]^2} + \ldots + \frac{[C_N+B_N]}{[1+i]^N}}\right\} + 2\left\{\frac{\frac{C_2}{(1+i)^2}}{\frac{C_1}{[1+i]} + \frac{C_2}{[1+i]^2} + \ldots + \frac{[C_N+B_N]}{[1+i]^N}}\right\}\right.$$
$$\left. + \ldots + N\left\{\frac{\frac{C_N+B_N}{[1+i]^N}}{\frac{C_1}{[1+i]} + \frac{C_2}{[1+i]^2} + \ldots + \frac{[C_N+B_N]}{[1+i]^N}}\right\}\right] \tag{4.11}$$

The numerator in each term represents a time of arrival, 1, 2, ..., N, of a payment that is weighted by the present value of that payment. In the denominator, we have the present value of the sum of all cash flows promised by the bond, which should be its current market price, $\bar{P}$. Define

$$W_t \equiv C_t/(1+i)^t \text{ for all } t = 1, 2, \ldots N-1 \tag{4.12}$$

as the coefficient attached to the payment to be received t years from now.[4] Let $w_N \equiv (C_N + B_N)/(1+i)^N$. Then, using (4.12) and the definition of P, we can write (4.11) as

$$\frac{dP}{P} = -\frac{di}{[1+i]}\left[\frac{(w_1 + 2w_2 + 3w_3 + \ldots + Nw_N)}{P}\right] \tag{4.13}$$

This equation gives the relationship between prices and yields. A fixed-income instrument's duration is its "price elasticity" and it relates percentage price changes to changes in yields. See Figure 4.8.

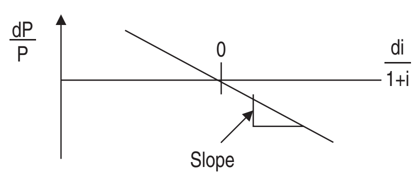


FIGURE 4.8 Duration

4. Each w_t is appropriately viewed as a "maturity coefficient" rather than a "weight" because the w_t's do not add up to one. However, each w_t divided by the denominator in (4.13) is a weight, that is, the $\hat{w}_t$s in (4.14) are weights.

Duration is the negative of the slope of the relationship shown in Figure 4.8. Thus, if we know the duration of an asset, we can predict its price sensitivity to a given change in yield. We can write:

$$\frac{dP}{P} = -D\left[\frac{di}{1+i}\right]$$

where D is duration. Defining $\hat{w}_t \equiv w_t/P$, we can write:

$$D = \sum_{t=1}^{N} t\hat{w}_t. \tag{4.14}$$

Thus, (4.14) says that, to arrive at the bond's duration, we compute a weighted average of the times to arrival of its different promised payments, where the weight attached to each time to arrival is equal to the present value of the cash flow associated with that time to arrival divided by the price of the bond.

We can think of the duration of a bond then as a metric for the average number of years a holder of that bond must wait before recouping his investment. For risk assessment purposes, duration is a much more meaningful attribute of a bond than its maturity. The shorter the duration of a bond, the lower is its price volatility. Holding everything else (including the current value or price of the bond) fixed, an increase in coupon payments reduces duration, and an increase in maturity increases duration. A zero-coupon (pure discount) bond has the longest duration among bonds of the same maturity; indeed, its duration is equal to its maturity. These bonds have recently become very popular. One significant advantage that they offer is that all cash flows they generate (which are only maturity) are implicitly reinvested at the YTM, rather than at the prevailing interest rate as with coupon bonds. However, zero-coupon bonds are also very risky because of their longer duration and consequent higher price volatility. When interest rates are falling, the holder of a zero-coupon bond realizes a greater price appreciation than the holder of an otherwise similar coupon-paying bond. But when interest rates rise, the holder of the zero-coupon bond also experiences a greater price decline! Let us see the effect of duration at work in the following simple illustration.

Duration at Work: Some Numerical Examples

The following key points about duration are worth noting:

1. Duration is denominated in years. It is a measure of the "weighted average life" of the bond.
2. Longer maturity assets have longer durations, *ceteris paribus*.
3. For zero-coupon bonds, duration = maturity. For all other bonds, duration < maturity. Holding everything else fixed, an increase in the coupon decreases duration.
4. The duration of a floating-rate instrument ("floaters") where the coupon changes with interest rates is the time until the next repricing.

Example 4.3 Consider an interest rate environment in which the one-period annual yield is 10 percent and the two-period annual yield is 9.7824 percent, and suppose we have two riskless bonds (each with a 2-year maturity) that are identical in all respects except that one is a zero-coupon bond that matures 2 years from now and promises a balloon payment of $1,109.60, where the other is a bond that will pay a coupon of $100 1 year from now and another coupon of $100 plus a balloon payment of $900 2 years from now. Compute the durations of these two bonds.

Solution We solve this problem in three steps. First, we compute the current prices of the zero-coupon bond and the coupon-paying bond using the yield data provided. We find that both are equally priced. Second, we calculate the duration of the coupon-paying bond, which is less than that of the zero-coupon bond. Finally, in step 3 we compute the variances of possible price changes (due to random interest rate movements) and show that this variance is higher for the zero-coupon bond.

Step 1 The discount rate for one period cash flows is 10 percent and the discount rate for two-period cash flows is 9.7824 percent. Thus the price of the zero-coupon bond is

$$P_0 = 1109.6/(1.097824)^2 = \$920.64.$$

Similarly, the price of the coupon bond is

$$P_c = \left[[100/1.10] + [1000/(1.097824)^2]\right] = \$920.64.$$

Step 2 The above calculation shows that both bonds are equally priced. The duration of the zero-coupon bond is its maturity, which is 2 years. The duration of the coupon-paying bond is

$$D = \hat{w}_1 + 2\hat{w}_1$$

where $\hat{w}_1 = [100/1.10]/920.64 = 0.09875$

and $\hat{w}_2 = [1000/(1.097284)^2]920.64 = 0.90125.$

That is, 9.875 percent of the value of this bond is attributable to its first period coupon and 90.125 percent of its value is attributable to the sum of its second period coupon and principal. Hence, $D = 0.09875 + 2(0.90125) = 1.90125$ years.

5. The duration of a bank's "core deposits" is typically taken as zero.
6. The duration of a portfolio is the weighted average of the durations of all the assets in the portfolio.

Using Duration to Measure the Impact of Interest Rate Shocks on a Bank's Equity Value: Recall that a bank's balance sheet can be expressed as

$$A = L + E$$

Where A = assets, L = liabilities and E = equity. Then, given a change in yield Δi, the balance sheet changes can be expressed as:

$$\Delta A = \Delta L + \Delta E \qquad [4.14]$$

Now:

$$\frac{\Delta A}{A} = -D_A\left[\frac{\Delta i}{1+i}\right]$$

which implies:

$$\Delta A = -D_A[A]\left[\frac{\Delta i}{1+i}\right] \qquad [4.15]$$

Similarly,

$$\frac{\Delta L}{L} = -D_L\left[\frac{\Delta i}{1+i}\right]$$

which implies

$$\Delta L = -D_L[L]\left[\frac{\Delta i}{1+i}\right] \qquad [4.16]$$

Assuming that the yield shock to the assets is identical to the yield shock to the liabilities, we can substitute (4.15) and (4.16) in (4.14) to obtain:

$$\Delta E = \left[-D_A[A]\frac{\Delta i}{1+i}\right] - \left[-D_L[L]\frac{\Delta i}{1+i}\right]$$

which implies:

$$\Delta E = \left[-D_A[A] + D_L[L]\frac{\Delta i}{1+i}\right]$$

or

$$\Delta E = -\left[D_A - D_L\left\{\frac{L}{A}\right\}\right][A]\left[\frac{\Delta i}{1+i}\right] \qquad [4.17]$$

where ΔE is in dollars.

So, when market yields change, what drives the change in the bank's equity value? There are three main drivers:

1) The size of the shock $\left(\frac{\Delta i}{1+i}\right)$
2) The amount of the leverage the bank uses
3) The mismatch between the durations of the bank's assets and liabilities. The bank will be "immunized" when $D_A = D_L \cdot L_A$

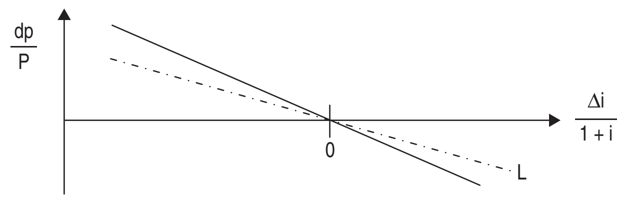


FIGURE 4.9 Asset and Liability Duration for Traditional Bank

How does this matter to a bank or a savings institution? To address this question, note that a traditional bank or savings institution has assets of longer duration than liabilities. Thus, its durations look like those shown in Figure 4.9.

What this means is that if yields increase, the bank's equity value declines (recall [4-17]), which shows that when $D_A > D_L$ and $L < A$, the term $\left[D_A - D_L \frac{L}{A}\right] > 0$, so $\Delta E < 0$ for any $\Delta i > 0$. If yields decrease, the bank's equity value increases. Thus, when a bank mismatches its balance sheet in the traditional way, it accepts interest rate risk in this way. Immunization closes the "gap."

A bank can alter its degree of immunization by changing the durations of its assets and liabilities. It can do this in two ways: on-balance sheet and off-balance sheet. On-balance sheet initiatives include making new types of loans, seeking new liabilities and changing its capital structure. Off-balance sheet initiatives include repurchase agreements, futures, options and swaps (we will discuss these in a later chapter).

Convexity

If a bank is interested in protecting its net worth against unexpected interest rate changes, duration matching can help; matching terms to maturity cannot do this unless all investments are of the zero-coupon variety. Suppose now that a bank is immunized and yields subsequently change. Does the bank remain immunized? The answer is no. The reason is that duration is an *approximation*. In fact, it is a *linear* approximation of a *nonlinear* relationship between prices and yields. We can see this with an example.

Example 4.4: Suppose we have a 10-year zero-coupon bond that is risk free, has a par value of $1,000, and is priced to yield 10 percent. What is its duration and how well will duration predict price changes if the yield moves up or down by 500 basis points?

Solution: Note that because this is a "zero" maturity = duration, so the duration here is 10 years. The current price of the bond is: $\frac{\$1{,}000}{(1.10)^{10}} = \$385.54$. Now consider the prices of this bond in response to a 500 basis point (b.p.) change in the yield.

Prices	**Yield Change**	
	+500 b.p.	−500 b.p.
Duration-Predicted Price: $\frac{\Delta P}{P} = -10\left[\frac{\pm 0.05}{1.05}\right] = \pm 47.62\%$	\$385.54[1 − 0.4762] =\$201.95	\$385.54[1 − 0.4762] =\$569.13
Actual Price	$\frac{\$1.000}{(1.15)^{10}}$ =\$247.18	$\frac{\$1.000}{(1.05)^{10}}$ =\$613.19
Error	−\$45.23	−\$44.78

We see then that duration *overpredicts* price declines when interest rates rise and *underpredicts* price increases when interest rates fall. Moreover, duration makes greater errors when yields rise than when they fall.

Why does duration make such prediction errors? The reason is that the true relationship between price changes and yield changes is *convex*, not linear.

When we first calculated the relationship between dP and di, we took a first derivative, which gave us the slope of the function in a "local" area, i.e., the slope of the curve, dP/di at di = 0. However, if we had gone further and computed the second derivative, we would have found $\frac{d^2P}{di^2} > 0$, i.e., all fixed-income securities are convex.

One implication of convexity is that duration will do a reasonable job in predicting price changes as long as interest rate changes are in the neighborhood of di = 0, i.e., relatively small changes like, say, 1 basis point. But the larger the interest rate change, the more erroneous duration is in predicting price changes. See Figure 4.10 below.

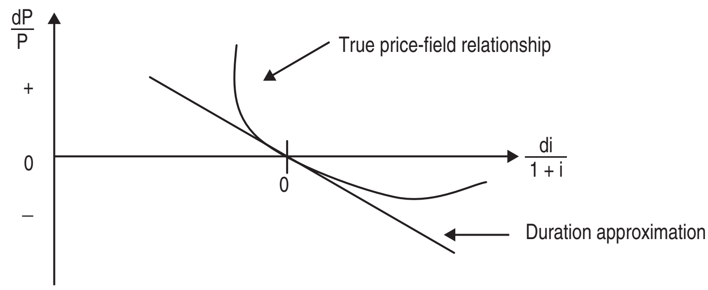


FIGURE 4.10 Price-Yield Relationship Is Convex

Implications of Convexity for Fixed-Income Securities and for Banks: There are three important implications of convexity for fixed-income securities:

1. The price decline given a rate increase is *smaller* than the price increase given a rate decrease of the same absolute magnitude as the rate increase.
2. Duration changes as yields change.
3. Greater convexity implies greater errors in the predictive ability of duration.

There are two important implications of convexity for banks:

1. Asset convexity is desirable. If the bank's asset portfolio is more than its liability portfolio, then properly done duration immunization never hurts the bank.
2. Duration immunization is a dynamic process since asset and liability durations change as yields change.

Interest Rate Risk

How Interest Rate Risk Can Affect a Financial Institution's Net Worth

The successful financial institution must understand its interest rate risk and manage the *durations* of its assets and liabilities. A *pure broker* need not worry about interest rate risk because its assets and liabilities are always duration matched. On the other hand, the *asset transformer* is often exposed to very subtle forms of interest rate risk. Consider the following simple example. A bank is borrowing and lending funds of two maturities: short term (1 year) and long term (2 years), all zero-coupon. Loans consist of \$40 million short term and \$40 million long term, while liabilities are \$60 million short term and \$10 million long term.[5] All numbers are in market value terms as of October 30, 2002. Hence, the bank's balance sheet is

Bank's Balance Sheet as of October 30, 2002

Short-term loans	\$ 40,000,000	Short-term liabilities	\$60,000,000
Long-term loans	\$ 40,000,000	Long-term liabilities	\$10,000,000
Total assets	\$ 80,000,000	Total liabilities	\$70,000,000
		Equity	\$10,000,000
		Total equity and liabilities	\$80,000,000

The yield curve as of October 30, 2002, is a flat solid line, as shown in Figure 4.11. Annual yields on assets and liabilities of all maturities are 10 percent.

Now suppose that on October 31, 2002, the yield curve shifts to the dotted line shown in Figure 4.11. All yields rise to 12 percent.

Each dollar of short-term assets (or liabilities) decreases in value to \$0.9821428 and each dollar of long-term assets (or liabilities) decreases in value to \$0.9646045. The new balance sheet in market value terms looks as follows

Balance Sheet as of October 31, 2002

Short-term loans	\$39,285,712	Short-term liabilities	\$58,928,568
Long-term loans	\$38,584,180	Long-term liabilities	\$ 9,646,046
Total assets	\$77,869,892	Total liabilities	\$68,574,613
		Equity	\$ 9,295,279
		Total equity and liabilities	\$77,869,892

Thus, the market value of equity falls by \$704,721 or 7.047 percent. The shift in the term structure affects the values of *both* the assets and the liabilities, but it has unequal effects on assets and liabilities due to unequal maturity weighting or duration. To see

5. You can easily verify that the asset and liability portfolios here have different durations.

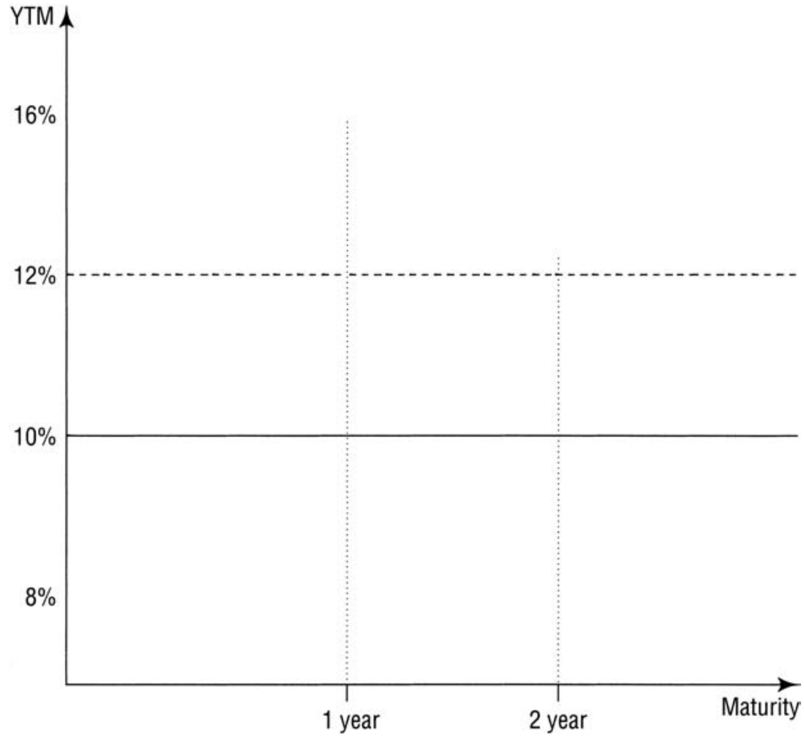


FIGURE 4.11 Yield Curves Facing Hypothetical Bank

this, note that the duration of short-term assets is 1 year and the duration of long-term assets is 2 years. The weights attached to the short-term and long-term assets are 0.5 and 0.5. Thus, the duration of the asset portfolio is $0.5 \times 1 + 0.5 \times 2 = 1.5$ years. Similarly, the duration of the short-term liability is 1 year and the weight attached to it is \$60 million/\$80 million $= 0.75$, while the duration of the long-term liability is 2 years and its weight is \$10 million/\$80 million $= 0.125$. Thus, the duration of the liability portfolio is $(0.75 \times 1 + 0.125 \times 2) = 1$ year.

While unequal duration weighting is risky, it is also a service provided by an asset transformer. By funding short (acquiring short-duration liabilities), the intermediary reduces the duration of its clientele's assets, thereby earning any term premium embedded in the yield curve. One simple way to eliminate interest rate risk altogether is to *equalize* the durations of assets and liabilities at all times. But then the institution forgoes duration/maturity transformation, a potentially profitable type of asset transformation.

A Case Study in Interest Rate Risk

Banks and other depository institutions often deliberately mismatch the durations of their asset and liability portfolios to profit either from term premiums or from their own expectations (guesses) about where interest rates are headed. Depository institutions characteristically fund their longer-lived assets with shorter-term liabilities. For instance, S&Ls historically funded 30-year fixed-rate mortgages with deposits that were often subject to withdrawal on demand.[6] Similarly, commercial banks would finance 5- and 7-year fixed-rate "term" loans with demand and savings deposits, both of which could be withdrawn at a moment's notice. Such mismatches inevitably entail interest rate risk.

As an illustration, consider NCNB Corporation, a North Carolina-based banking company, which later went on to become Nations Bank and then merged with Bank of America; the bank speculated that there would be an interest rate downturn in 1990.[7] It thus lengthened the duration of its investment portfolio through 1989. At year end, the bank had a liability-sensitive balance sheet, largely because of its holdings of $6 billion in long-term Government National Mortgage Association (GNMA) mortgage-backed securities. As of December 31, 1989, about $1.5 billion more of NCNB's liabilities than its assets would have repriced over the next 12 months. If interest rates had fallen, NCNB would have enjoyed a huge profit. Instead, interest rates rose. As of year end 1989, the yield on 30-year GNMAs was 9.49 percent. By March 16, 1990, the 30-year GNMA yield was 9.95 percent. NCNB consequently suffered a $180 million unrealized loss in its bond portfolio. That news—plus disclosures in March 1990 that problem loans could rise by 25 percent in the first quarter of 1990—sent NCNB's stock plummeting from $46 in the first week of March 1990 to $40 by March 19, 1990, a decline of 12 percent.[8]

The Savings and Loan Experience and Other Episodes

Another striking example of the consequences of interest rate risk is the experience of the U.S. savings and loan (S&L) industry in the 1980s. S&Ls have traditionally financed themselves with short-maturity deposits and invested in relatively long-maturity, fixed-rate mortgages. Consequently, their liabilities repriced more frequently than their assets. As long as the yield curve sloped upward, this was a profitable maturity transformation. But in the late 1970s and early 1980s, the yield curve inverted as yields rose to historic highs. S&Ls took significant losses. This dissipation of much of the industry's net worth was the triggering event that led to the decimation of the industry years later. In particular, the loss of net worth meant that these institutions had much to gain and little to lose by pursuing risky investments. This led to further losses. The financial distress of Orange County in California in the 1990s is another example of the potentially devastating effect of interest rate risk.

Why Take On Interest Rate Risk?

The immediate question is: Why do banks and S&Ls *choose* to accept such exposure? That is, we have seen that it is possible for the bank to avoid taking much of the interest rate risk it normally takes on simply by matching the durations of its assets

6. Because of mortgage prepayments, 30-year fixed-rate mortgages have uncertain duration, typically of 7 to 12 years.

7. This discussion was reported by Kelly Holland in *American Banker*, March 20, 1990.

8. As Mr. John W. Munce, senior vice president and balance-sheet-management executive at NCNB put it, "We were postured to benefit from falling rates over a 12-month horizon. We definitely took some losses."

and liabilities, so interest rate risk is largely an avoidable risk. To answer this question, we go back to the theory of the term structure of interest rates. The presence of a risk premium in the term structure invites those who are more risk tolerant than the "average" (or representative) investor to hold long-term assets and fund these assets with shorter-term liabilities. Their reward is the premium in the yield curve that reflects the greater risk aversion of the respective investor.

Why should banks and other depository institutions be more risk tolerant than others? This is an issue we will take up in later chapters, but for now it suffices to note that deposit insurance may be one reason for a preference for risk on the bank's part. Of course, not all banks will desire to take on the same amount of risk. As in the case of their borrowers, the risk-taking propensities of banks depend on their own capital levels. Banks with more capital may wish to make investments that are less risky than those desired by banks with less capital.

The upshot of this discussion is not that an asset transformer should not take interest rate risk, but rather that such risk must be carefully assessed and managed.

Liquidity Risk

Our discussion of liquidity risk proceeds as follows. First, we introduce the concept of liquidity risk and discuss what liquidity risk means for a bank. We then present some formal definitions of liquidity. This follows with discussions of ways in which depository institutions manage liquidity risk. Finally, we end with a discussion of how a central-bank-based solution to the liquidity problems of individual depository institutions creates a moral hazard of its own.

What, After All, Is Liquidity Risk?

There are occasions on which the bank does not have ready access to funds that it needs, and is therefore forced to incur costs. These could be the costs associated with passing up investment opportunities. Alternatively, they could be distress financing costs. These are examples of situations in which the financial intermediary faces liquidity risk. We define *liquidity risk* as the risk of being unable to satisfy claims without impairment to its financial or reputational capital.[9]

Informational frictions are at the heart of liquidity problems. To see how informational asymmetries interact with default and interest rate risks to create liquidity risk, let us imagine that you own a bank that has made loans of $1 million with a maturity of 2 years and financed them with uninsured demand deposits. As a banker, you know more about the default risk of your loans than outsiders do, that is, there is asymmetric information about loan quality. Now, suppose that 6 months down the

9. It is important to distinguish between illiquidity and insolvency. The latter relates to a condition in which the value of the firm's liabilities exceeds the value of its assets, and hence its net worth is negative. Illiquidity can be as damaging and costly as insolvency, but it is a form of distress rooted in the (non)marketability of assets rather than in their ultimate or full value. To be sure, this may be a vacuous distinction when addressed at close range. Nevertheless, in thin markets, time and marketing efforts often are essential to the realization of asset values. Liquidating assets on short notice often results in "distress" prices. The relationship between time available for marketing and the realizable values of assets is central to the notion of liquidity.

road, \$400,000 of deposits are withdrawn, but your existing stock of cash assets is only \$100,000. This means you need to raise \$300,000 to fund the deposit withdrawal. If potential depositors' perceptions about the quality of your loan portfolio are sufficiently favorable, you will not have any trouble acquitting new deposits in the amount of \$300,000. But suppose that outsiders have received unfavorable information about your loans.[10] If this information is sufficiently unfavorable, new deposits may simply not be forthcoming,[11] or you might have to pay an excessively high interest rate—relative to the rate you consider "appropriate"—to attract the necessary deposits.[12] This is an example of liquidity risk.

There are two points we should note about this example. First, an informational asymmetry about asset quality plays a pivotal role in creating liquidity risk. If outsiders knew as much about your loan quality as you do, then you would be able to acquire the deposits you need at a price that you consider appropriate for the risk associated with the loan portfolio. This eliminates liquidity risk.

Second, duration mismatching may be an important ingredient in creating liquidity risk, but it is not a necessary ingredient. To see the importance of duration mismatching, suppose your asset and liability portfolios were perfectly duration matched. Then the assets that were funded by a specific set of liabilities would pay off at the same time that the liabilities came due, and informational asymmetry about these assets that arises *after* these assets are on the bank's books would not matter. Of course, if an informational asymmetry exists about the new loans you make, then a premium reflecting this asymmetry will show up in the interest rate on the deposits raised to fund these loans. However, you can pass this premium along to your borrowers in the way you price your loans, so that your capital is not impaired.

The Interaction Between Liquidity and Default Risks

However, you could have liquidity risk even with a duration-matched balance sheet. If some of the loans funded by deposits were to default, then withdrawals of these deposits would need to be funded in part by new deposits, assuming that loan defaults are large enough to leave insufficient liquidity to finance the withdrawals. Unless you plan to make new productive investments, depositors would have little reason to provide new deposits. Thus, new deposits would not be available *just* to finance old deposit withdrawals. To see this in the context of the previous example, suppose that both loans and deposits have a 2-year maturity. However, due to loan defaults, only \$1 million is collected from loan repayments at maturity, whereas deposit withdrawals at maturity amount to \$1.3 million. New deposits of \$300,000 must be raised to finance withdrawals. This amount can only be raised against new assets that you acquire. Suppose now you wish to make \$2 million in new loans with a 2-year maturity and thus need to raise \$2.3 million in new deposits (ignore equity capital for now) that will also have a 2-year maturity. If your assessment of the quality (repayment probability) of these loans is higher than that of depositors in general, then the deposit interest rate will exceed what you believe is justified by the default risk of

10. This information may be different from what you know about your loans, that is, you may still know more than outsiders and may thus believe that your loan quality is good.

11. Indeed, it is possible that all of your existing deposits may be withdrawn.

12. In fact, your willingness to pay such a high rate of interest may be viewed as a signal of poor loan quality. Then, liquidity risk can be interpreted as the likelihood of incurring this signaling cost.

your loans. Suppose that, in present value terms, the excess amount you must pay in deposit interest is $46,000, that is, 2 percent of the total deposits acquired.

How much of this excess amount can you pass along to your borrowers? The answer depends on competition in the credit market. For simplicity, suppose that any other lender would face the same problem in communicating information about these loans to potential depositors, that is, any lender would suffer a cost equivalent 2 percent of the total deposits. However, a bank that does not need to finance *old* deposit withdrawals would need to raise only $2 million in deposits. Hence, 2 percent of $2 million can be passed along to borrowers in the form of a higher loan interest rate.[13]

Returning to your bank, then, if you are to be competitive in pricing your loan portfolio, you'll be able to pass along $40,000 in excess deposit interest to your borrowers. But that means you are stuck with a $6,000 "out of pocket" expense that arises because of your lack of sufficient liquidity to meet the excess of deposit withdrawals over net loan revenues. The possible incidence of such a cost is part of liquidity risk. Note again that this cost arises because there is a problem of asymmetric information about your loans. With perfect information, liquidity risk is not an issue here.

The Interaction Between Liquidity and Interest Rate Risks

We now turn to the interaction between interest rate risk and liquidity risk. There are two ways to explain this interaction. First, suppose we have deposit interest rate ceilings. Given this ceiling, a rise in market interest rates causes withdrawals because depositors can earn higher rates elsewhere. Hence, deposit interest rate ceilings transform interest rate risk into withdrawal risk.

Another way to understand this interaction is by returning to the example we discussed in the section under interest rate risk. If the term structure receives a random shock that causes interest rates to rise, it is possible that you will experience a deposit outflow as your depositors will want to reinvest their money at the prevailing higher interest rates. You have two ways to finance these withdrawals. One way is for you to acquire new (partially insured) deposits. But this may require you to pay a premium to depositors due to a possible informational asymmetry about your loan portfolio. Moreover, you must satisfy reserve and capital requirements on deposits. An alternative is to liquidate part of your asset portfolio to meet these unanticipated deposit withdrawals. You can do this by selling off marketable securities you hold or by selling off some of your loans.[14] Due to an informational asymmetry about your loans, however, you may only be able to sell your loans for less than what you think they are worth. The loss you incur as a result is also a part of liquidity risk. Although this loss is precipitated by an unfavorable move in interest rates, note again the

13. The assumption here is that there are many competing banks that can make the loans in question, and each of these banks needs to raise $2 million in deposits to finance $2 million in loans.

14. A bank can sell its loans to another bank just as a firm would sell its debt in a private placement. This practice, which is quite old, is known as "loan sales." A more recent practice is securitization, which involves the bank selling the loan, typically as a component in a portfolio of loans, directly to investors in the capital market. This is usually done through an underwriter and is a process of converting a previously untraded security into a traded security. We will have a lot more to say about this in Chapter 9.

central role played by asymmetric information. Moreover, the greater the asymmetric information, the greater the potential for loss, and hence the lower the asset's liquidity. This is why, despite an active secondary market, a corporation's common stock is not as liquid as a U.S. Treasury bill.

Some Formal Definitions of Liquidity

Think of P^* as the full-value price of an asset, that is, the highest price an owner can expect to realize by liquidating one unit, provided all useful preparations are made for the sale. If the asset is sold before all useful preparations can be made, a lesser price will be realized. Call this lesser price P_i, where $i = 0, \ldots, n$ indicates the time used for marketing, and n is the time needed to realize full value. The length of time used should be thought of as the interval between a decision to sell and the time at which a sales contract is consummated.[15] Hence

$$P_n = P^*$$

and for all values of $i < n$, the realized price of the asset, P_i, is less than full value. One way to think of liquidity is in terms of

$$L_1 = \frac{P_i}{P^*}.$$

A limitation of this definition is that the liquidity of a particular asset depends on the value of i chosen. Thus, for low values of i, one asset may be more liquid than another, whereas for greater values of i, the liquidity comparison might be reversed. This impedes the consistent ranking of assets according to their liquidity. One way to mitigate, if not obviate, this problem of liquidity reversal among assets is to think in terms of an "average" value of i. Hence

$$L_2 = \sum_{i=0}^{n} \frac{P_i}{P^*}.$$

A still more appealing approach recognizes the inherent uncertainty regarding i, the time interval between the decision to sell and the actual sale. Thus, we can view it as a random variable with a probability distribution, g(i), which stipulates the probability of each possible outcome ($i = 0, \ldots, n$). The expected value of an asset, E(P), is then defined as

$$E(P) = \sum_{i=0}^{n} g(i)P_i,$$

15. The terms of the transaction are fixed at the time the sales contract is consummated, but the transfer of property takes place at the "closing," a date that may coincide with the date of the sales contract, but often occurs later.

and this leads to a third definition of liquidity, which is

$$L_3 = \frac{E(P)}{P^*}.$$

The liquidity concept can be further generalized to account for marketing expenditures, say M. The more general view is that the realizable price of an asset depends on time, marketing expenditures, and full-value price, so that

$$P_i = f(i, M, P^*),$$

and if $\overline{M}$ is the optimally chosen marketing expenditure,

$$E(P') = \sum_{i=0}^{n} g(i) f\left(i, \overline{M}, P^*\right)$$

is the expected value of an asset, conditional on the owner's spending optimally on marketing. This leads to our fourth definition of liquidity

$$L_4 = \frac{E(P')}{P^*}.$$

and M/P^* can be thought of as a measure of the market's thinness, a measure akin to the bid-ask spread.[16]

Note that the positive relationship between available time for marketing and marketing effort on the one hand and realizable value on the other has nothing to do with changes in supply or demand for the asset; the realizable value increases in the context of given market conditions. Time is not used to await a more favorable market, but rather to do the marketing necessitated by costly information. For a depository institution, there are many ways to reduce liquidity risk. An obvious way is to simply keep more liquid assets on hand. The other is to reduce the deposit withdrawal risk that creates liquidity risk. A third way is to rely on a lender of last resort who stands ready to replenish the bank's liquidity when needed. In what follows, we discuss each in turn.

Reducing Liquidity Risk With Liquid Assets

Think of the fractional reserve banking system described in Chapter 3. That bank can be thought of as holding two kinds of assets: cash and loans that mature in two or more periods (prior to maturity the loans are assumed to be worthless). The bank's liabilities all mature in one period, and may or may not be renewed (withdrawn). If the fraction withdrawn after one period is equal to, or smaller than, the bank's holding of cash assets, the bank will continue in business for two periods, at least.

16. For a fuller development of this idea, see Greenbaum (1971).

On the other hand, if withdrawals exceed the bank's holding of cash assets, that bank will be unable to honor its liabilities—it has promised all depositors immediate access even though its own capacity to satisfy claims is strictly limited by its holding cash assets.[17] Therein lies the liquidity conundrum of banking.

Notice that an important role of a bank is the provision of liquidity services, and it provides this service by mismatching its balance sheet on the liquidity attribute, that is, it holds assets that are less liquid than its liabilities. This is one form of asset transformation. The *quality* of this liquidity service provided by the bank depends on three factors: the liquidity of its loan portfolio, the cash (or liquid assets) it has on hand, and the withdrawal risk in its deposit base. By investing in more liquid loans and/or keeping more cash on hand, the bank can improve its own liquidity. However, it does so at the expense of profits. An alternative would be to seek ways to dissipate withdrawal risk, which is what we turn to next.

Reducing Liquidity Risk by Dissipating Withdrawal Risk

A depository institution can reduce the variance of its deposit flows by diversifying the sources of funding, that is, having many distinct and dissimilar depositors. This is formally demonstrated in Appendix 4.2. A diverse depositor base results in more predictable deposit flows; the improved predictability reduces the cash needed to service a deposit base to any arbitrary probabilistic standard. That is, the larger and more diverse the depositor base, the smaller the cash holding necessary to achieve any preselected probability of a stock-out (liquidity crisis). This is one way the depository institution *produces* liquidity. Nevertheless, withdrawals will sometimes exceed the institution's capacity to service them, even though this may happen only with very small probability, and in that sense the system is imperfect. Indeed, this is the system's Achilles' heel. Bank runs are the trauma that illustrate this vulnerability of fractional reserve banking, a vulnerability caused by the illiquidity of bank assets.

Reducing the Liquidity Risk of an Individual Bank With a Lender of Last Resort

It was long ago discovered that the liquidity of a fractional reserve banking system can be ensured with a thoroughly credible "lender of last resort" (LLR). This was the major motivation for the creation of central banks, including the Federal Reserve System. With an institution capable of creating money limitlessly, it becomes possible to support banks facing the most extraordinary deposit outflows. Provided that the banks are sound (solvent, given reasonable time to liquidate their assets), this could

17. This is the rationale behind the standard measure of liquidity in the savings industry, which is the ratio of cash and short-term U.S. government securities and other specified securities to deposits and borrowing due within 1 year. The Office of Thrift Supervision (OTS) has established minimum liquidity requirements for savings institutions.

be done by having the central bank lend to the banks using their illiquid loans as collateral. With such a lending facility, sound but illiquid banks could be protected and financial market disruptions avoided. This argument is developed more fully in Appendix 4.2.

However, an inexpensive, readily available LLR faces the danger of inheriting the entire liquidity management problem of the banking industry. That is, the bank's incentive to hold cash assets (or even diversify its deposit base) is weakened if borrowings from the central bank are inexpensive and readily available. This is a moral hazard associated with the introduction of the LLR, and it has two implications. First, it shifts deposit seigniorage from the public to privately owned banks. Second, the LLR is also exposed to the credit risk of the bank's collateral. The moral hazard of lower, voluntarily held cash assets explains the consequent introduction of cash asset reserve requirements, and also why there are carefully administered detailed rules and informal restrictions governing access to the discount window.

Thus, legal reserve requirements and LLR pricing and availability shift at least a portion of the liquidity management problem back to the banks. Other banks, without access to an LLR facility, own the liquidity problem outright.

Closing Remarks on Liquidity

The management of liquidity is referred to as the treasury function, and it is usually entrusted to the chief financial officer (CFO). It is her responsibility to "fund the bank." This requires a professional understanding of the institution's cash flows, as well as all potential sources of liquidity. Ultimately, protection comes from maintaining diverse, capacious, and reliable sources of funding against future contingencies. This explains why the typical bank will borrow from virtually *all* reasonably priced sources. To be sure, cost will be a consideration, but opportunities to reduce short-run funding costs by concentrating on fewer funding sources are commonly avoided.

In "paying up" for funding diversity, the bank is purchasing lines of credit, and this reduces the likelihood of being rationed. It is common for funding sources to evaporate under stress; CFOs understand this only too well. Continental Illinois Bank and Trust found that holders of its large CDs (Certificates of Deposit) abandoned them in their hour of keenest need, and the high-yield bond market went into eclipse when Drexel Burnham Lambert was forced into insolvency because banks chose to withdraw their funding. The conventional protection against the trauma of being rationed is to accept the extra cost of participating in as many markets as possible, thereby diversifying funding sources. Liquidity is consciously purchased by banks as well as their borrowers, and it is the fragility of liquidity that makes this part of banking particularly challenging.

Conclusion

Like any other firm, a bank faces risks that can be managed but not totally avoided. For a bank, the three major risks are default risk, interest rate risk, and liquidity risk. These are interrelated and their interaction depends in an important way on the presence of asymmetric information. The current approach is to manage these three risks and others holistically as part of Enterprise Risk Management (ERM).[18]

Interest rate risk is linked to the term structure of interest rates. Our analysis of the term structure both under certainty and uncertainty shows how yield and maturity are related. In both the certainty and uncertainty cases, the concept of riskless arbitrage plays a key role. Further, our analysis shows that the risk in holding a bond is more appropriately assessed in terms of its duration rather than its term to maturity. The definition of duration and the examination of its relevance in measuring the price volatility of bonds indicate how coupon-paying bonds should be analyzed. We also examined the concept of convexity and measures of interest rate risk exposure. We followed this with an examination of liquidity risk and the interaction of liquidity and interest rate risks. With these tools in hand, we considered the management of these risks by the bank.

Case Study Eggleston State Bank

Introduction

Mr. Edward Eggleston, CEO and primary stockholder of Eggleston State Bank, the bank he founded some 30 years ago in his hometown of Bloomington, OR, is worried. He has just gotten off the phone with an old friend of his, Fred Fisher. Fred had reported the difficulties he was having with his job search.

Fred's and Edward's life stories were remarkably similar. College roommates, they had both founded small hometown banks in the years following college and had managed to be quite successful for a number of years. But now, Fred is effectively wiped out—his bank has been closed by regulators and his fortune, invested entirely in the bank, has evaporated. Currently, he is going through the process of looking for a new job, maybe in the sort of big city he had always prided himself on avoiding.

Fred's bank had been fairly small, with $30 million in total assets, but had been consistently profitable as a small-town bank doing traditional banking—accepting deposits from individuals and small businesses in the short-term, while making long-term mortgage loans and business loans. But when state banking regulations were relaxed, allowing a branch of a major state bank to move into town, things got tighter. This competition, along with increasing volatility in interest rates and the bank's traditional mismatching of its balance sheet, led the bank into a situation with increasingly deteriorating capital, with a drop in capital over a 3-year period from $2 million to under $300,000. Finally, regulators moved in and took over the bank.

Edward Eggleston sighs, and wonders to himself whether the same thing could happen to his bank. His bank is much larger than Fred's with total assets of over $400 million (see Exhibit A). But with the rise of several regional banks with assets in billions of dollars, Edward is beginning to feel like he may face the same kinds of problems that beset Fred's bank, in the form of increased competition from larger, more sophisticated banks. He decides to meet with his executives to carefully

18. See, for example, Nocco and Stulz (2006).

investigate the exposure of Eggleston State Bank to interest rate risk, and to discuss the possibilities for hedging against changes in interest rates.

The Meeting

A week later, Edward Eggleston is sitting in his office with Carol Chipley and Douglas Date. Carol is a recent graduate of a top MBA program with strong analytical skills, hired in part to help modernize the bank's approach to risk management. Douglas, on the other hand, has risen to his current position from within the bank, primarily due to his sharp eye for detail and sound common sense.

Eggleston: O.K., gang. You both know our situation as well as I do. What I'm interested in is what options we have for action, and which you think we ought to pursue. Should we remain mismatched, or is it time for us to move into hedging?

Date: Well, as you know, Ed, I've always been skeptical about us getting involved in the latest fads in banking. After all, I don't see that we are so mismatched. Remember that article I showed you a while back, about a bank that started fooling around in the futures markets on the bad advice of a smooth-talking broker? I'm afraid that if we aren't careful, we could wind up making a big mistake. Besides, we've been here for 30 years now, steadily profitable. Why should we mess with a good system?

Chipley: I think that you are right, Douglas, when you say that we should be careful. But I think that for every story about banks losing money because a hedging program was poorly planned, we can find a dozen stories about banks that lost money, or even went under, because they weren't hedged at all. Plus, the banking environment has changed significantly in recent years. So what worked for the last 30 years might be fatal to us over the next 30 years.

Date: People are always saying that, but I don't really see what has changed. We've gotten bigger, but this is still a small-town bank. Our borrowers and our customers are mostly individuals and small to medium-sized businesses. Carol, weren't you just showing me the other day a chart showing how smooth our deposit flows have been over the past 5 years? (See Exhibit B). And the new administration seems committed to keeping Ft. Washington open, so it looks like the overall business outlook for the community is about the same as it ever was: stable and solid as a rock. This is a fairly prosperous area, after all. (See Exhibit C).

Chipley: Well, I'm not so sure that we can count on any administration keeping promises about military bases. But anyway, closing Fort Washington isn't the only risk that we face. I think that the increasingly competitive nature of banking means that world markets can affect what happens in our little town. Twenty years ago, our customers might not have worried so much about differences in interest rates; we were their hometown bank and we knew them and their business. But banking is more impersonal now, and we can't just expect our depositors to stay with us if we don't offer competitive interest rates. I think our investment and loan portfolios deserve a careful look (see Exhibits D and E).

Eggleston: Well, those are the reactions that I expected to hear from you. But I think that now is the time for some hard-boiled analysis. Let's sit down right now and come up with some likely interest rate scenarios. Then Carol can work with the figures and let us know exactly what would happen to the bank under a variety of circumstances (see Exhibit F).

The Numbers

Exhibit A
EGGLESTON STATE BANK
Year-End Balance Sheets (in Thousands)

	2004	2005
Assets		
Cash & due from banks	$59,696	78,645
U.S. govt. obligation	$38,612	45,284
Other govt. obligations	$58,030	49,456
Other securities	$6,678	6,439
Loans and discounts	$250,950	290,125
Bank premises	$12,698	21,924
Other assets	$2,996	2,876
Total assets	$429,660	494,749
Liabilities		
Demand deposits	$178,668	184,694
Time deposits	$122,164	166,995
Deposits of the U.S. govt.	$10,164	3,429
Other govt. deposits	$57,190	59,805
Due to commercial banks	$7,266	12,987
Total deposits	$375,452	427,910
Other liabilities	$23,520	34,925
Total liabilities	$398,972	462,835
Capital Accounts		
Common stock	$5,838	5,630
Capital surplus	$15,008	14,472
Undivided profits	$7,952	9,828
Reserves	$1,890	1,985
Total capital accounts	$30,688	31,915
Total liabilities and capital accounts	$429,660	$494,750

Exhibit B
Total Deposits (in Millions of Dollars)
(Expected Duration Six Months)

	High	Low	Daily Average
2001	305	257	284
2002	323	291	301
2003	363	323	357
2004	375	307	363
2005	427	375	400

Exhibit C
Market Area Economic Data
Income and Housing

Annual Household Income	Percentage of Households
Under $3,000	28%
$3,000–$6,999	20%
$7,000–$14,999	30%
$15,000–$24,999	21.5%
$25,000+	.5%
Home Ownership	
All Housing Units	30,000
Owner-occupied	51%
Rental	38%
Unoccupied	11%
Major Area Employers, Bloomington	
Ft. Washington	25,000
Lockheed	1,000
Kraft Foods	850
Bloomington College	730

Exhibit D
Eggleston State Bank
(Investment Portfolio, Today)

Description	Par Value	Coupon	Years to Maturity	Book Value	Bond Rating
U.S Government Securities					
Bills	2,500,000	—	8 months	2,235,000	—
Notes	4,000,000	6.00	2 years	3,765,000	—
Bonds	40,000,000	7.00	25 years	39,284,000	—
Other Government Securities					
Municipal Securities	50,000,000	6.00	22 years	49,456	Baa
Corporate Bonds					
Lockheed	7,000,000	12	17 Years	6,439	Aaa

Exhibit E
Eggleston State Bank
(Loan Portfolio, Summary Report, Today)

Borrower Type	Coupon	Estimated	Book Value
Short-Term Individual (Cars, and so on)	13.27	2.1	14,700,000
Short-Term Business	12.31	1.8	7,234,000
Medium-Term Business	11.45	5.3	42,300,000
Long-Term Business	10.4	7.9	78,766,000
Home Mortgages	8.3	9.1	179,000,000

During the meeting, the bankers came to an agreement on the following probabilities for the following scenarios:

Exhibit F

Likely Interest Rate Scenarios
(Scenario Names)

	Good	Bad	Ugly
Probability	.5	.3	.2
U.S. Govt. Securities			
Bills	11.00%	9.00%	12.00%
Notes	10.00%	10.00%	13.00%
Bonds	9.00%	11.00%	14.00%
Other Govt. Securities			
Municipal Securities	9.25%	11.75%	15.25%
Corporate Bonds			
Lockheed	9.75%	10.75%	13.75%
Loans			
Short-Term Individual	13.25%	11.25%	14.25%
Short-Term Business	12.25%	10.25%	13.25%
Medium-Term Business	10.50%	10.5%	13.75%
Long-Term Business	9.80%	10.75%	13.75%
Home Mortgages	9.00%	11.50%	14.50%

The Assignment

Eggleston: Carol, I'd like for you to take these numbers and report back to me on some very specific questions. What exactly is the extent of our mismatching? What would happen to the bank under the various scenarios that we've talked about? What kind of hedging program, if any, should we use to protect the bank?

Review Questions

1. What are the three major types of risks faced by banks?
2. What is the term structure of interest rates?
3. Under certainty, if the term structure is determined to preclude riskless arbitrage, what is the relationship between the yields on bonds of different maturities and why?
4. What is duration and why is it a more valid metric to consider for coupon-paying bonds than maturity? What is the relation between duration and price volatility for bonds with the same maturity?
5. What is convexity? Discuss its potential usefulness in evaluating bonds.
6. Discuss the pros and cons of duration mismatching for a depository institution.

7. What is liquidity risk and how is it linked to interest rate and credit risks? What is the role of asymmetric information in creating liquidity risk?
8. How can liquidity risk be managed? What are some of the impediments faced by banks in implementing an *integrated* risk management system that manages credit risk, liquidity risk, and interest rate risk?
9. Suppose there are three zero-coupon bonds, identical in all respects except maturity. Each bond has a face value of $1,000. One of them matures a year from now and is currently selling at $855.66. Another matures 2 years form now and is currently selling at $835.33. The third matures 3 years from now and is currently selling at $775.85. Compute the YTM for each of the three bonds, plot the yield curve (assuming that you can interpolate smoothly), and compute the available forward rates.
10. The annualized YTM on a single-period pure discount bond is 12 percent and that on a two-period pure discount bond is 10.45 percent. There are two bonds. One is a two-period, pure discount bond that promises a balloon payment of $1,200 at maturity. The other is a bond that will pay a coupon of $100 one period hence, and a coupon of $100 plus a balloon payment of $1,000 two periods hence. Compute the duration of these bonds and their possible price changes prior to maturity.
11. Given below is an excerpt from "A Friendly Conversation." Provide a critique.

Moderator: So, what do you people think? Will we ever really understand what happened to the American banking industry well enough to know what should be done?

Appleton: Well, I think banks and S&Ls were simply victims of the environment. We had an inverted yield curve—long rates were lower than short rates—for a while and this made it difficult for financial institutions to reap their normal profits from asset transformation; you know, I've never believed in the expectations hypothesis. It's a theoretical nicety with no practical relevance. Of course, the increased interest rate volatility didn't help. As if this wasn't enough, there was an enormous increase in competition, both domestic and international. These institutions must have felt like they were being squeezed by a powerful vise.

Moderator: By the way, Alex, I'll give you another reason not to like the expectations hypothesis—it's also wrong.

Appleton: I didn't know that. Are you sure? In any case, it's good to know you agree with me, Mike. But frankly, I'm surprised. Knowing how you and Beth feel about this, I thought I'd get more of an argument.

Moderator: Well, cheer up, Alex. My agreement with you is only partial. I agree that depository financial institutions faced a tough environment during the last 15 years or so. But I also think they could have *managed* their risks more intelligently. For example, they could have reduced the duration gaps in their asset and liability portfolios and made use of contemporary immunization techniques to hedge their interest rate risks. Like some of the investment banking houses, they could have been more innovative in brokerage activities, so that the resulting fee income would have made banks less dependent on the riskier asset transformation activities. Just look at the profits earned by some investment bankers who stripped Treasuries and sold zeros (pure discount bonds) like CATS (Certificates of Accrual of Treasury Securities) and TIGRS (Treasury Investment Growth Receipts). No, Alex! The real story runs much deeper than your "passive victims of the environment" explanation. I think banks and S&Ls *exploited* the system and ripped off taxpayers.

Appendix 4.1 Dissipation of Withdrawal Risk Through Diversification

Suppose that a bank has n depositors, each of whom deposits \$1. Each deposit is subject to withdrawal after one period, but may remain for two. Assume that the probability that a \$1 deposit will be withdrawn after one period is one in ten, that is, $p = 0.1$, but whether a given deposit is actually withdrawn after one period cannot be known until that one period has passed.

Deposits are used to fund loans that pay back in full in two periods, but are worthless until they mature. (There is no secondary market in loans.) This is a harmless simplifying assumption and does not affect the argument that follows. Of course, the bank will need to hold some fraction of its assets in cash in order to satisfy its one-period withdrawals. The question is how much cash the bank should prudently hold. If the bank has \$1 or \$1 million of deposits, the probability of withdrawal remains fixed at 10 percent, and the expected withdrawal is this probability multiplied by the amount of deposits. However, if the bank has only \$1 in deposits, the withdrawal inevitably will be all or nothing at all, zero or one. Indeed, the expected value of \$0.10 is unattainable, and the bank's decision to hold 10 percent in cash, if feasible, is virtually pointless.

However, as the bank's depositors increase in number, assuming independence among them, the withdrawal of 10 percent becomes more predictable; in the limit, as depositors become more and more numerous, a 10 percent cash holding will "almost certainly" satisfy deposit withdrawals.

This idea is apparent from the definition of the standard deviation of a binomial distribution where n is the number of depositors and $q \equiv 1 - p$; the standard deviation of the bank's deposits will be $\sigma = \sqrt{npq}$.

Note that this measure of uncertainty varies with the square root of the *number of depositors*, and hence in the limit as the number of depositors increases to infinity, the standard deviation per dollar of deposit equals $\lim_{n \to \infty} (\sigma/n) = 0$.

This means that as the number of depositors becomes larger, the withdrawal uncertainty *per loan* diminishes, approaching zero in the limit, even though the withdrawal probability remains unchanged at $p = 0.1$. So, as the depositor population increases, the 10 percent withdrawal can be treated increasingly as a routine (almost fixed) cost, rather than as a potential catastrophe. The risk of ruin, the probability that withdrawals exceed the bank's cash holding, never actually becomes zero since $\sigma/n \to 0$ only in the limit. But the risk of ruin can be managed, and made indefinitely small by diversifying the bank's sources of funding.

Appendix 4.2 Lender-of-Last-Resort Moral Hazard

In a world of fiat money, value derives from an administered or artificial scarcity. That is, our money is money by fiat or legal mandate (hence legal tender) and is not convertible into gold or any other commodity at a fixed exchange rate, as in the case of commodity-backed money. The more money the government prints, or otherwise creates, the less its value, and this applies to bank deposits as well as to paper money. The administered scarcity of money also creates a monopoly profit referred to as "seigniorage." This profit on the production of money is shared by the privately owned banks and the public, via its effective ownership of the central bank. The Federal Reserve is nominally owned by member commercial banks. However, the equity in the Federal Reserve banks pays a statutorily fixed rate of return, much like a bond, whereas the residual earnings of the Federal Reserve flow back to the U.S. Treasury via a special franchise tax. Given that neither central bank nor private bank deposits pay interest (any interest rates below competitive rates will sustain the point), the distribution of seigniorage between the banks and the public (or central bank) depends on the cash asset reserves the banks choose to hold. The more reserves banks hold, the smaller will be banks' share of the seigniorage.

Since the introduction of an LLR reduces the amount of reserves the banks will desire to hold, it effectively shifts seigniorage from the public to the banks. This is the moral hazard associated with the introduction of an LLR, and it explains that one rationale for legal reserve requirements (that stipulate the minimum cash assets that banks must hold) is to restore the "appropriate" sharing of seigniorage between banks and the public.

This point is easily illustrated. Suppose we have a single commercial bank with \$10 million in deposit liabilities, an amount consistent with the money supply the central bank wishes to maintain in consideration of monetary policy. There are no reserve requirements and no LLR facility. The commercial bank voluntarily holds 10 percent of its assets in cash against withdrawal risk. It makes no difference whether the bank's cash assets are vault cash or deposits at the central bank, so for simplicity assume these assets are all on deposit at the Federal Reserve where they earn nothing. The commercial bank's balance sheet would then be

Commercial Bank

Cash assets	\$1 million	Deposit liability
Loans or other earning assets	\$9 million	\$10 million
Total assets	\$10 million	Total liabilities \$10 million

The Federal Reserve's balance sheet, to a first approximation, would show

Federal Reserve

Earning assets	\$1 million	Deposit liability \$1 million

Note that the Federal Reserve's deposit liability corresponds to the bank's cash assets. Now suppose the Federal Reserve introduces an LLR facility. It has no reason to change the money supply, but banks now have a new source of liquidity. Hence, they will feel less need to hold nonearning cash assets. Say they cut these holdings from 10 to 5 percent. The bank's balance sheet now becomes

Commercial Bank

Cash assets	\$0.5 million	Deposit liability
Loans or other earning assets	\$9.5 million	\$10 million
Total assets	\$10 million	Total liabilities \$10 million

and the Federal Reserve shrinks to

Federal Reserve

Earning assets	\$0.5 million	Deposit liabilities	\$0.5 million

In effect, \$0.5 million in earning assets have been transferred from the Federal Reserve's balance sheet to the bank's balance sheet, and this occurs as a direct consequence of the introduction of the LLR.

One could argue that if the LLR facility is properly priced, the moral hazard will be discouraged. However, note that before its introduction, the LLR interest rate was infinite, so that any finite interest rate will improve bank liquidity, and should therefore result in some reserve dissipation. As a historical matter, the LLR tends to price low for reasons that are not entirely clear. This generous pricing practice aggravates the moral hazard problem and heightens the need for legal reserve requirements.

Thus, reserve requirements control the moral hazard of the LLR, and a lowering of reserve requirements transfers deposit seigniorage from the public to the banks. Raising reserve requirements has the reverse effect. One hundred percent reserve requirements shift all deposit seigniorage to the public. This is the basis for the conventional wisdom that the reserve requirement is a tax on the banks, but one could just as easily argue that any reserve requirement less than 100 percent is a subsidy to banks. The hard question here is: To whom should the monopoly rents associated with administered money belong?

References

Cox, John, Jonathan Ingersoll, and Stephen Ross, "A Theory of the Term Structure of Interest Rates," *Econometrica* 53–2, March 1985, 385–407.

Fisher, Irving, and Roman L. Weil, "Coping with the Risk of Interest Rate Fluctuations," Journal of Business 44–4, January 1971, 408–431.

Greenbaum, Stuart I., "Liquidity and Reversibility," *Southern Economic Journal* 38–1, July 1971, 83–85.

Ho, Thomas S.Y., and Sang-bin Lee, "Term Structure Movements and Pricing Interest Rate Contingent Claims," *Journal of Finance* 41–5, December 1986, 1011–1030.

Holland, Kelley, "Capital: NCNB Loses Big Bet on Long-Term Rates," *American Banker*, March 20, 1990, 20.

Ingersoll, Jonathan E., *Theory of Financial Decision Making*, Rowman and Littlefield, New Jersey, 1987.

Ingersoll, Jonathan E., Jeffrey Skelton, and Roman L. Weil, "Duration Forty Years Later," *Journal of Financial and Quantitative Analysis* 13–4, November 1978, 627–650.

Klotz, Richard G., "Convexity of Fixed Income Securities," manuscript, updated.
Macaulay, F., *The Movements of Interest Rates, Bond Yields, and Stock Prices in the United States Since 1856*, New York: National Bureau of Economic Research, 1938.
Nocco, Brian W., and René M. Stultz, "Enterprise Risk Management: Theory and Practice," working paper, Ohio State University, July 2006.
Sprenkle, Case, and Merton H. Miller, "The Precautionary Demand for Narrow and Broad Money," *Economica*, November 1980, 407–422.

This page intentionally left blank

CHAPTER • 5

Spot Lending

"Neither a borrower nor a lender be; For loan oft loses itself and a friend, and borrowing dulls the edge of husbandry."

William Shakespeare

Glossary of Terms

Loan: The extension of credit via a typically untraded and illiquid debt contract.

Security: A financial claim, debt, or equity, which may be traded or untraded.

COD: Cash on delivery as a method of payment for goods received.

Commercial Paper: Unsecured debt, offered as a short-maturity (less than 270 days) security by corporations.

T-bills, T-notes, and T-bonds: Debt securities of varying maturities issued by the U.S. government through the U.S. Treasury Department; hence, "T" for Treasury.

FHLB: Federal Home Loan Bank. The Federal Home Loan Bank System, headed by the Federal Home Loan Bank Board, was formerly the primary regulatory agency for savings and loan associations. The district home loan banks are now providers of financial services, including liquidity, to smaller commercial banks and thrifts.

FHLMC: This stands for the Federal Home Loan Mortgage Corporation. Also known as "Freddie Mac," its basic function is to facilitate the provision of liquidity to lenders by purchasing existing mortgages from their portfolios. It finances these purchases by borrowing from the Federal Home Loan Banks, issuing

GNMA-guaranteed mortgage-backed bonds, selling mortgage participation certificates on which it guarantees interest and principal, and selling guaranteed mortgage certificates.

FNMA: This stands for the Federal National Mortgage Association. It is a privately owned (stockholder-owned), government-sponsored enterprise. Also known as "Fannie Mae," its basic function is to provide a secondary market in trading and securitizing home mortgages. It is the largest purchaser of residential mortgages in the United States. Its activities are similar to those of Freddie Mac, except that it faces no statutory limitations on the organizations with which it can conduct business.

GNMA: This stands for the Government National Mortgage Association. This is a wholly owned, corporate instrumentality of the U.S. government, operating within the Department of Housing and Urban Development (HUD). Also referred to as "Ginnie Mae," its role is to enhance liquidity in the market for mortgages. Ginnie Mae does this in a variety of ways. For example, many mortgages carry a fixed interest rate so that when market interest rates rise, existing mortgages sell at a discount (that is, at less than face value). Ginnie Mae issues a commitment to the mortgage seller (the originating financial institution, for example) to purchase the mortgage at a fixed price. After acquiring the loan, Ginnie Mae sells it to "Fannie Mae" at the prevailing market price. Ginnie Mae absorbs any discount from the price paid to the seller. Another function of Ginnie Mae is to guarantee securities backed by government-insured or guaranteed mortgages. That is, Ginnie Mae provides guarantees for securitized claims against portfolios of government-insured mortgages.

S&P Stock Index: Standard & Poor's composite index of 500 large-company stocks.

Incentive Compatibility: A condition that requires the alignment of incentives between the agent and the principal. See Chapter 1.

C&I Loans: Commercial and industrial loans. These are loans extended to nonfinancial firms.

Nash Equilibrium: A steady state attained when none of the contracting parties has an incentive to change its actions unilaterally. See Chapter 1.

HLT: Highly leveraged transaction, which is a loan to a borrower with a very high debt/equity ratio.

Collateral: An asset used to secure a loan. Failure to repay the loan completely and in time transfers the collateral to the lender.

Absolute Priority Rule: A rule that prioritizes creditors' claims to a borrower's assets according to their seniorities.

GAAP: Generally Accepted Accounting Principles.

Prime Rate: A reference/benchmark borrowing rate posted by the bank for its better customers.

LIBOR: London Interbank Offer Rate. This is the interest rate banks charge each other for short-term loans in the United Kingdom.

CD Rate: The interest rate offered by banks on certificates of deposit.

Optimal Stopping Rule: A statistical decision rule that tells the decision-maker when to stop a sequential sampling process and make a decision. For example, a bank may have $1 million to lend and knows that the longer it waits, the more loan applicants it can screen before deciding who to lend the money to. However, waiting is costly because of the time value of money. An optimal stopping rule in this case would specify conditions under which the bank would find it most profitable to stop screening further loan applications. Another example is determining when a bank should stop acquiring additional information about a borrower, and make a decision.

Discriminant Analysis: A statistical technique used to identify the factors most useful in predicting an event. An example would be the factors useful in predicting bankruptcy.

The Glass-Steagall Act: An act passed by Congress in 1933 to separate commercial and investment banking in the United States. It prohibits commercial banks from engaging in securities underwriting and other investment banking activities as well as the activities of insurance companies.

Introduction

For many commercial bankers, lending is the heart of the business. Loans dominate asset holdings and account for a large share of revenues and costs. Lending takes place in both spot and forward credit markets. We begin here with a discussion of spot lending.

The purpose of this chapter is to explore the asset side of the bank's balance sheet. We begin in the next section with a brief review of the most prominent assets on a bank's balance sheet. The following section explains what we mean by lending, and the difference between loans and securities. We also discuss how these assets are purchased. The structure of loan agreements is discussed in a subsequent section. This is followed by a section that discusses the major informational problems in loan contracts and the importance of (perceived) loan performance for the determination of a bank's stock price. The next section examines credit analysis. Our emphasis is on the economic underpinnings of the various traditional factors considered in credit analysis. In particular, we relate these economic underpinnings to the informational problems pervasive in loan contracting. In the section that follows, we turn to sources of credit information. We consider both internal sources within the bank and external sources such as financial information agencies. In the next section, we take up analysis of borrower's financial statements. We follow it up with a section on the examination of loan covenants. Our focus is on the *why* of each covenant. A case study follows the concluding section.

Description of Bank Assets

Trends in the Composition of Bank Assets

There are three basic types of assets on a bank's balance sheet: loans, marketable securities, and cash. See Figure 5.1. Before we discuss each of these in detail, we will briefly review recent trends in the composition of bank asset portfolios.

In Figure 5.2 we show the time-series behavior of the composition of commercial bank assets. While loans have risen slightly as a fraction of total assets in the late 1970s

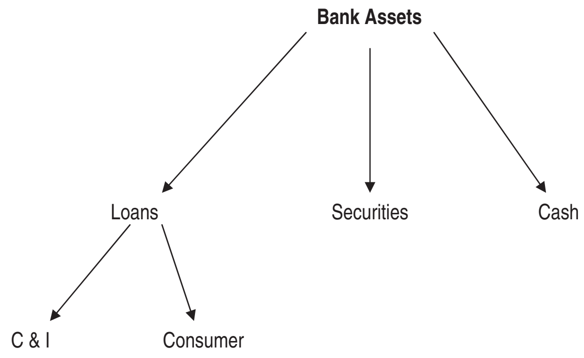


FIGURE 5.1 Spot Lending

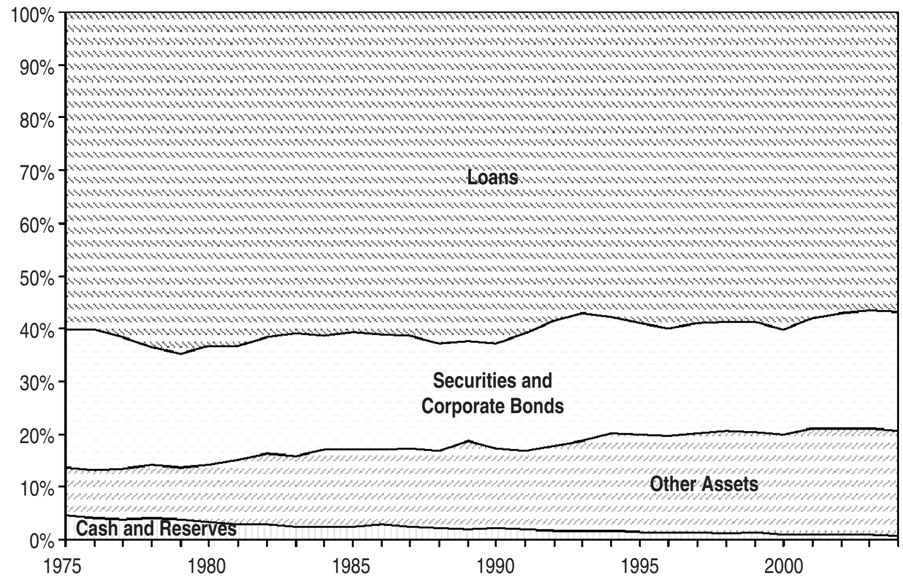


FIGURE 5.2 Composition of Commercial Bank Assets
Source: Federal Reserve Statistical Release: Flow of Funds Accounts of the U.S. 1975–1984, 1985–1994, and 1995–2004.

and 1980s, they declined slightly thereafter. Security holdings declined slightly in the late 1970s and have been relatively steady since. Cash and reserves have declined quite a bit, and this decline has been consistent through time. A clearer picture of what has been going on emerges from Figure 5.3, which shows the time-series behavior of commercial bank loans. It is apparent that C&I loans have declined in relative importance as banks have increased their mortgage holdings. Consumer credit has declined slightly in percentage terms from 20 percent in 1975 to 15 percent in 2004.[1]

1. Consumer loans are mainly comprised of credit cards, installment loans, mortgages, and home equity loans. These are essentially "commodity products," with apparently little product differentiation across banks. However, they still leave open considerable room for product innovation. For example, Wells Fargo gained prominence in the consumer loan market with its hybrid of a fixed-rate mortgage and an adjustable-rate loan. Moreover, the effectiveness with which credit information is processed is crucial in determining the attributes of consumers to whom these loans are made, and hence their profitability.

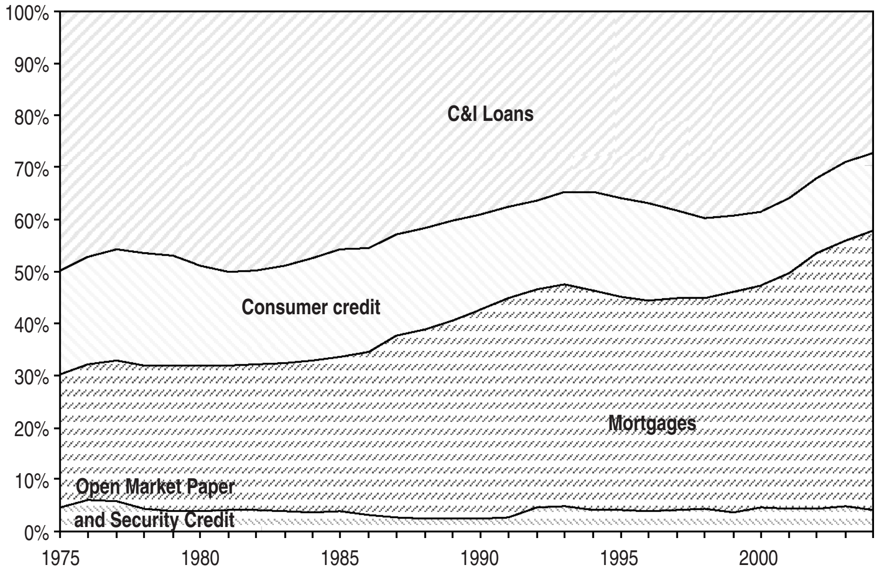


FIGURE 5.3 Composition of Commercial Bank Loans
Source: Federal Reserve Statistical Release: Flow of Funds Accounts of the U.S. 1975–1984, 1985–1994, and 1995–2004.

There are two main reasons for this trend. The first has to do with the changing nature of commercial lending. A bank has an advantage over the capital market in providing credit to a firm as long as banks have cheaper access to loanable funds than investors, *and/or* banks can resolve private information and moral hazard problems more effectively. Over the years, much of the deposit-related rents available to banks have eroded, thereby extinguishing virtually all of the funding advantage possessed by banks. Moreover, with the boom in financial innovation in the last two decades, a variety of new securities have been used by firms to raise funds directly from the capital market. These securities, as well as the securitization of bank-originated loans (see Chapter 9), have been designed to cope with the very problems of private information and moral hazard that banks have specialized in solving.[2] Thus, the relative advantage of banks over the capital market in providing credit to firms has diminished. With the capital market becoming a more viable source of competition in the commercial lending arena, the profitability of lending to large corporations has declined significantly for banks; hence, the relative decline in C&I lending.[3]

Types of Bank Loans

We will first discuss business loans, often referred to as C&I loans, which fall into four main categories.

2. For example, Green (1984) shows that a convertible bond (that is, a bond that can later be converted to stock by the bondholder) can be effective in controlling the moral hazard problem stemming from the borrowing firm's inclination to invest in risky projects to the detriment of bondholders.

3. For example, Security Pacific acquired $2.7 billion in mortgages in mid-1990 when it successfully bid for Gibraltar Savings, the largest California thrift under government control.

(a) **Transaction Loans:** A transaction loan is negotiated for a specific purchase and is tailored to the particular needs of the purchaser. The demand for these loans from a particular borrower is typically episodic and hence each loan is negotiated separately. The loan is usually secured by the asset being financed with the loan (for example, equity in another company), and repayment is expected to come from the use of this asset.

(b) **Working Capital Loans:** These loans are used by firms to finance routine day-to-day transactions. Thus, they are general purpose, short-term borrowings and are often used either to purchase current assets (like inventories) or to repay debts incurred in purchasing current assets. These loans are also usually secured by collateral such as accounts receivables or inventories.

(c) **Term Loans:** These are longer maturity loans used to buy fixed assets requiring large outlays of capital. Maturities typically run from 3 to 10 years. Repayment is normally amortized because it comes out of the cash flows generated by the asset financed with the loan. Borrowings are almost always drawn down under revolving lines of credit or similar commitments.

(d) **Combinations:** Working capital loans often include provisions that permit the conversion of short-term borrowings into term loans at the borrower's request.

We will now briefly review consumer loans.

(a) **Consumer Loans (excluding mortgage loans):** The most important types of consumer loans are direct loans and bank credit card receivables. A *direct consumer loan* is typically financing for the purchase of durable goods such as cars, boats, or appliances, and is secured with the asset being purchased. *Bank credit card* borrowings are a form of short-term, unsecured general purpose credit. Credit cards became widely used in the mid-1960s. Credit card lending has proved to be very profitable for banks.[4] The profitability of bank credit cards stems from three sources: (i) the discount at which the bank purchases sales slips from merchants (this discount typically ranges from 2 percent to 6 percent), (ii) the interest rate charged to a card user who chooses not to remain current in payments (most cards extend an interest-free grace period based on a monthly billing cycle), and (iii) the annual membership fees charged to credit card users.[5]

(b) **Mortgage Loans:** These are a specialized form of consumer and commercial lending. The purpose of a mortgage loan is to finance the acquisition or improvement of real estate. These loans are almost always secured by the real estate they finance. The three principal types of mortgage loans are: residential mortgage loans, construction loans, and commercial mortgage loans.

Until the advent of securitization, mortgage loans were illiquid assets because of the uniqueness of each property, the severity of private information problems, and the uncertain maturity of the loan due to the possibility of prepayment by the borrower. However, securitization took care of many of these impediments to the marketability of mortgages and facilitated the liquification of these instruments. This

4. See Ausubel (1990).
5. Many banks waive these annual fees because of increased competition for credit card business.

was especially true in the market for residential mortgages where Fannie Mae and Freddie Mac led the way under government auspices.

Securitization is a technology for transforming illiquid loans into traded liquid securities by separating the origination of the instrument from its funding. Typically, a financial institution such as a bank "originates" the loan, that is, it screens the applicant, designs the loan contract, and determines the pricing parameters. However, instead of using deposits to fund the loan as in the traditional case, the bank sells the loan to a special trust that assembles a portfolio of loans and funds the portfolio in the capital market, often with the advice and assistance of an investment banker. The services provided by the investment banker include the sale of claims against the loan portfolio to investors and then the maintenance of a secondary market in the securitized claims. The enormous growth in securitization in the past two to three decades is evidence of its benefits in the mortgage market. These benefits stemmed from the liquidity created by the standardization, diversification, possible subsidies provided by the government via Fannie Mae and Freddie Mac, and new contract design that accompanied securitization. Securitization is discussed in greater detail in Chapter 9.

Earlier, fixed-rate mortgages—in which the borrower's interest rate is fixed over the life of the mortgage—dominated the market. However, since the legalization of adjustable-rate mortgages (ARMs) in the 1980s there has been an explosion in the variety of mortgage designs. The terms of mortgages are as varied as the needs of borrowers and the imagination of lenders.

Marketable Securities Held by Banks

(a) Bankers Acceptances: These instruments arise mostly in connection with international trade. A bankers acceptance is a bank-guaranteed indebtedness of the bank's customer to a third party. This instrument usually arises as a time draft written by a firm in order to pay for some goods either in local currency or in foreign exchange. The draft is then "accepted" by the bank, that is, the bank guarantees its face value at maturity. The acceptance is then either held by the bank or sold in the secondary market and may be held by another bank. The originating bank typically charges a fee for the guarantee (acceptance) that is independent of the interest paid on the borrowing. Maturity is usually less than 6 months.

A bankers acceptance facilitates trade between parties that operate in different legal systems with wide geographical and cultural separation. If the exporter does not know the importer well enough, it will not ship goods, even on a COD basis. However, it is likely that the importer's bank is better known and hence its willingness to guarantee payment—which serves the purpose of substituting its own credit risk for that of the importer—facilitates trade. The bank issuing the guarantee also can be expected to know more about the importer, usually a customer of the bank. Its informational advantage vis-á-vis the exporter allows the bank to earn a fee on the acceptance. Thus, bankers acceptances are closely tied to the bank's role in providing a more efficient resolution of informational problems. For more on this, see Chapter 8.

(b) Commercial Paper: This is unsecured debt issued on the strength of the issuer's name. It is sold on a discounted basis like Treasury bills,[6] with maturities ranging

6. There is no explicitly stated interest rate, but the claim is sold at a price less than its face value (value at maturity), the difference implicitly defining the interest cost. Note, however, that discount yields are not directly comparable to bond yields; a translation is required to achieve comparability.

from 3 to 270 days and interest rates typically lower than prime and comparable to those on CDs and bankers acceptances. Only the best-known firms issue commercial paper because it is sold *directly* to investors, without an intermediary to resolve informational problems.

(c) U.S. Government Securities: These are important instruments for commercial banks because of their default-free nature and the highly liquid markets in which they are traded. As we saw in Chapter 3, private information content undermines liquidity, so U.S. government securities—which embody virtually no private information—provide banks with liquidity.

Income from all U.S. government securities is subject to federal income taxes as well as capital gains tax, but is exempt from state and local income taxes. Marketable U.S. government securities are of three types: bills, notes, and bonds. Treasury bills (T-bills) are short-term U.S. government securities (with original maturities of 91 days, 182 days, and 1 year) that, like commercial paper, are sold on a discounted basis. Treasury notes are similar to T-bills except that they have maturities not less than 1 year and not more than 7 years. Treasury bonds are issued with original maturities that often exceed 10 years, and can be as long as 30 years.

(d) U.S. Government Agency Securities: These are certificates of indebtedness issued by agencies of the U.S. government, such as the Federal Intermediate Credit Bank, the Federal National Mortgage Association (FNMA or Fannie Mae), the Federal Home Loan Bank (FHLB), and the Government National Mortgage Association (GNMA or Ginnie Mae). They are not direct obligations of the U.S. government, and they typically trade at a small premium over Treasury debt. Income on these securities, like direct U.S. government obligations, is exempt from state and local taxes, but not from federal taxes.

(e) State and Local Securities and Municipal Bonds: These debt instruments usually have a higher after-tax yield than Treasury and agency securities of comparable duration because of higher default risk and weaker liquidity. Their interest payments are exempt from federal income taxes as well as from home-state and local taxes. State and local government bonds can be divided into three broad categories: housing authority bonds, general obligation bonds, and revenue bonds. Housing authority bonds are issued by local housing agencies to build and administer housing. They are guaranteed by the federal government and are therefore virtually riskless. A bond is called a general obligation bond if the full faith and credit of the issuer stands behind the debt. In contrast, the interest and principal of a revenue bond is supported solely by the cash flow of a designated public project or undertaking. The revenues supporting these bonds may come from: (i) specifically dedicated taxes such as those on cigarettes, gasoline, and beer, (ii) tolls for roads, bridges, and airports, (iii) rent payments on buildings, office spaces, and the like. Typically, the bond payments are linked to the revenues produced by the project the bonds were used to finance.

(f) Other Assets: These include vault cash and deposits at the Federal Reserve, equity in subsidiaries, physical capital like buildings, computers, and loans originated by other banks that may have been acquired by the bank as part of a loan sale or through securitization. For short periods of time, the bank may also possess a variety of other assets acquired as collateral from delinquent borrowers.

What Is Lending?

A Definition

What is a *bank loan*? Simply put, it is the purchase of an asset (the borrower's indebtedness) that is typically an illiquid and highly customized financial claim against the borrower's future cash flows. In effect, the bank is obtaining from the borrower the legal right to a prespecified portion of the borrower's future cash flows over a prespecified period of time, and paying the borrower the present value of these cash flows. The bank's claim represents the borrower's repayment obligation and the loan amount represents the present value of these future obligations, assuming no extraordinary profit for the bank.

Methods of Acquiring Loans

There are two principal methods by which banks acquire loans: through *spot market purchases* and through *forward market purchases*. In the spot market, the bank can either originate the loan and then fund it by keeping the loan on its own books, or it can purchase the loan from another intermediary that originated it. A spot loan is created when the bank extends credit to a loan applicant immediately upon approval of the application. In the forward market, the bank issues a *promise* to the applicant that it will lend in the future on prespecified terms. Such a promise is known as a *loan commitment*. The bank commits to lend to the borrower up to a certain amount in the future on terms that are prespecified and at the option of the borrower. In this case, the bank is committing to purchase a financial claim from a particular borrower at some time in the future.

We discuss these two methods of asset acquisition in separate chapters. Spot market purchases and forward lending are covered in separate subsequent chapters. This division is merely for expositional convenience. In practice, the volume of spot and forward lending are inextricably linked. The extent of spot lending by the bank depends on how many of its outstanding loan commitments sold in previous periods are exercised or taken down in the current period. In general, a higher volume of takedowns of outstanding loan commitments implies a lower volume of spot lending in the current period, although the *total* volume of lending in the current period may rise (relative to that in the previous period) because of an unexpectedly high take-down on previously made commitments. This follows from the size constraints on banks associated with financial and human capital limitations.

The Decomposition of the Lending Function

The Decomposition: The subtlety of lending transactions is often blurred in the bundling together of distinct services relating to credit transactions. The normal commercial bank loan is logically decomposable into origination (the broker), funding (the lender), servicing (the collector), and risk-processing services (the guarantor). And lending can be thought of largely as credit risk management that includes these four activities as well as the bank's credit culture. See Figure 5.4.

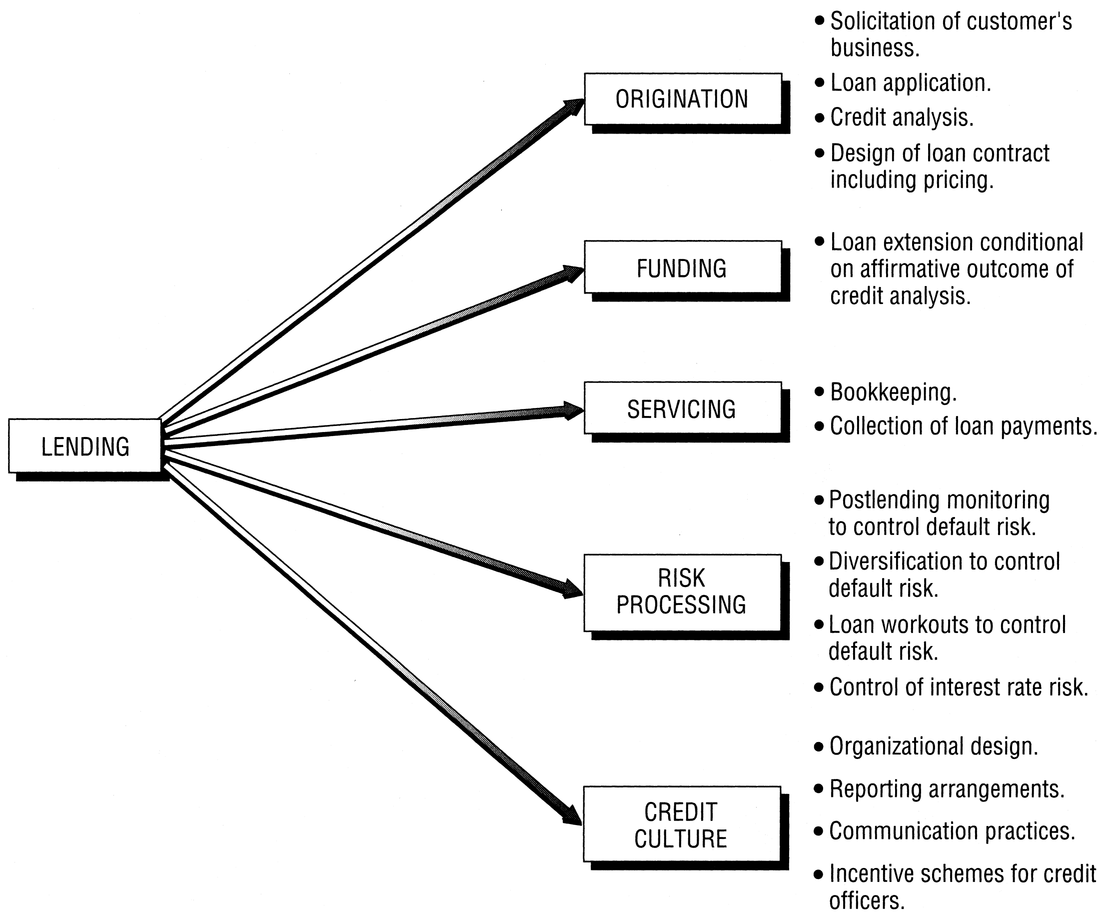


FIGURE 5.4 Decomposition of the Lending Function

Origination involves the activity of initiating a loan to a borrower. It is often described as the initial solicitation of the borrower and the screening of the loan application by the bank. Origination includes credit analysis and the design of the loan contract, both of which we discuss at length later in this chapter. *Funding* is the actual extension of the loan after an affirmative decision is reached in the credit analysis process. *Servicing* involves collecting loan repayments and keeping records. *Risk processing* involves postlending monitoring to control default risk, as well as activities designed to control the bank's interest rate risk arising from a loan duration that differs from the duration of the bank's liabilities. The *credit culture* involves the bank's organizational design, reporting arrangements, communication practices, and incentive schemes for credit officers. We will discuss credit culture later in the book. Much of our focus in this chapter and the next will be on origination (in particular, loan contract design and credit analysis) and risk processing (in particular, the control of default risk).

Industry Specialization: In the thrift industry (savings and loan associations, mutual savings banks, and the like) different institutions provide distinct credit services, which is clear evidence of institution specialization. For example, the mortgage banker originates loans and the mortgage processor services the loans. The loan is typically funded by the public (the net saver or surplus spending unit) in the form of newly purchased savings and loan deposits or mortgage-backed securities. The bulk of the credit and interest rate risk is sustained by savings and loan stockholders, the U.S. government (FDIC, FSLIC, NCUA, GNMA), specialized private insurers (for example, Mortgage Guarantee Insurance Corporation), or some combination of the three (FNMA, FHLMC). In commercial banking, it is common for the bank to hold originated loans. Consequently, the origination, servicing, and risk absorption is evidenced by an earning asset on the bank's balance sheet. The bank depositor holding a risk-free asset is funding the bank loans. The government, through the FDIC and the bank stockholders, shares the risks (uninsured depositors may sustain some exposure as well). Should the bank sell a loan, say to a closed-end mutual fund (as in the case of a savings and loan association selling mortgages to FHLMC or Salomon Brothers for packaging into a mortgage-backed security), then the security holder would do the funding and the location of the risk would depend on the specific terms (recourse or nonrecourse) of the sale. Irrespective of the terms of the sale, however, the bank need show no earning asset on its balance sheet and virtually all the same services would have been performed and the same exposures sustained without any accounting evidence thereof. This statement requires some qualification in that if a loan is sold with recourse, the accountant will probably insist on booking the asset, but if a loan is sold without recourse and a letter of credit is issued insuring against default (the above are equivalent), the balance sheet will show no loan and the letter of credit will probably appear in a footnote to the balance sheet, but not in its body.

In fact, banking reserve and deposit insurance premiums provide banks with an incentive to sell, rather than hold, earning assets. In this way, the bank can avoid these costs.

The traditional subsidy inherent in deposits (owing to underpriced deposit insurance, Regulation Q, and entry restrictions) encouraged banks to *hold* earning assets whereas deposit insurance premiums, reserve and capital requirements, along with less explicit regulatory costs, were a partial offset to the deposit subsidy. However, the deposit subsidy is rapidly disappearing, whereas many of the regulatory costs remain. Thus, we can predict that banks will de-emphasize the holding of the loans they originate, service, and guarantee. The recent emphasis on "fee income" is a reflection of this phenomenon.

Loans Versus Securities

In the previous discussions, we have talked about *loans* and *securities* as two distinct claims. The way we have defined loans, there is little difference between loans and debt securities, except that the latter are usually more liquid. That is, securities are traded in secondary markets, whereas loans usually are not. Loans are essentially *private* debt placements with banks. You will recall from our discussions in Chapter 4 that liquidity and marketability are interrelated. From an economic viewpoint, the distinction between loans and securities is in their *relative liquidity*.[7]

Viewed in this light, recent developments in the loan market can be seen as narrowing the distinction between loans and securities. We refer to *loan sales* and *securitization*. A loan sale, which is a fairly old practice, is simply the selling of a loan by the originating bank to an alternative funding agent, usually another bank. This can either be an *outright sale* of the loan, where the loan may have been originated by a single bank or as part of a *loan syndication*. With an outright sale, the originating bank disengages itself from the loan, that is, it makes the initial loan and then turns around and sells it to another bank, thereby removing the loan from its own balance sheet. A fee is earned for the originating service, so that the transaction leaves its mark on the originator's income statement. With an outright loan sale, the bank acts as a pure broker, although in practice almost every loan sale involves the originating bank retaining a part of the loan, so the bank is not a pure broker. Some loans are also made under syndication arrangements in which case there is *joint* origination of the loans by several banks. These loans may then be sold to others. Again, the "lead banks" in the syndicate earn fees.

Thus, loan sales enhance loan liquidity, especially if the originator maintains a secondary market. This blurs the distinction between loans and securities. A more recent practice for improving loan liquidity is securitization, which we discuss in detail in Chapter 9. Both loan sales and securitization trivialize the distinctions between loans and securities.

Structure of Loan Agreements

Trends in Loan Agreements: Commercial bank lending was once a fairly simple business. Most business loans were short-term, self-liquidating working capital credits, and terms were often left to informal agreements between a bank and its customers. Business lending began getting more complex in the 1930s when banks started making loans with maturities of more than a year, so-called term loans. Relations between banks and business borrowers have been growing more complex—and more formal—ever since.

Part of the push for more formality and variety in the design of agreements comes from the need for banks and borrowers to protect themselves from movements in interest rates over the credit cycle. Increases in market interest rates boost the costs to banks of funding outstanding loans and also reduce the attractiveness of existing credits. Reductions in market interest rates, on the other hand, often trigger prepayments.

7. From a legal standpoint, however, the distinction between a loan and a security was crucial during the Glass-Steagall Act, which prohibited commercial banks from engaging directly in securities activities. The statutory definition of a "security" is an expansive one; see Huber (1989). According to the 1934 Securities Exchange Act, the term "security" means not only any stock, bond, debenture, and evidence of indebtedness, but also the "countless and variable schemes devised by those who seek the use of the money of others on the promise of profits." However, a general exception is made for situations where the *context* makes it inappropriate to treat an instrument as a security. For example, a loan participation purchased by a depository institution from another institution is not considered a security. The minimum consequence of concluding that an instrument is a security is that the antifraud provisions of the securities laws become applicable. In practice, therefore, the distinction between a loan and a security is driven largely by legal interpretation that cannot always be supported on economic grounds. With the dismantling of the Glass-Steagall Act, this distinction has become somewhat of a moot point.

Floating interest rates have been one of the most important innovations in bank lending since the advent of the term loan. Provisions for adjusting loan rates periodically give banks some protection against interest rate risk. By combining the advantages of term and short-term loans, floating rates have allowed banks to compete for a share of the business credit market—even in the face of increased competition from the commercial paper market and other nonbank credit suppliers. At the same time, floating rates have effected changes in the other terms and conditions of commercial lending. An unintended consequence has been the loss of some borrowers who switched from banks to the capital market to obtain longer-term debt with greater fixity in the borrowing rate.

Details of Loan Agreements: A loan agreement specifies the obligations of borrower and lender, makes certain warranties, and usually places certain controls and restrictions on the borrower. It states the amount to be borrowed, or *the principal*. The agreement also states *the maturity*: short-term (less than 1 year), intermediate-term (1 to 5 years), and long-term (greater than 5 years). *The pricing formula* also is stated. The interest rate may be a fixed or a floating rate. If the interest rate is floating, it may be "prime-plus" (for example, the prime rate plus 1 percent) or "times-prime" (for example, the prime rate times 1.05). Pricing might also be at a "transaction rate," that is, the bank agrees *ex ante* to a fixed mark-up over a current money rate (for example, T-bill, the negotiable CD rate, or the commercial paper rate). The agreement also states the closing fees to be paid when the loan gets funded. In a competitive situation this fee may be 0.25 percent to 0.375 percent, and higher in other situations. Also, a penalty or default rate of interest may be stipulated for late or early payments.

Although loan agreements usually are tailored to meet the requirements of specific situations, most contain certain standard provisions, which may be divided into three general categories: conditions precedent, warranties (also called representations), and covenants and events of default.

The *"conditions precedent"* section includes requirements the borrower must satisfy before the bank is legally obliged to fund the loan. These conditions may include specific business transactions that must be completed or events that must have occurred. Other standard items are the opinions of counsel, certificate of no defaults, the note, and resolutions of the borrower's board of directors authorizing the transaction.

The *"warranties"* section of the loan agreement contains information and assumptions about the borrower's legal status and creditworthiness. By executing the loan agreement, the borrower attests to the accuracy and truth of the information provided as of the date of execution. Misrepresentation constitutes an event of default. Principal warranties include the following:

- A warranty that all financial statements submitted to the lender are genuine and fairly represent the financial position of the borrower (that is, that no material adverse change has occurred).
- The borrower has a valid title to all assets.
- The borrower has complied with all federal, state, and municipal laws and is not involved in litigation.
- The borrower has filed all necessary tax returns and has paid all taxes due.
- No need for third-party consent.
- No violation of existing agreements.
- Collateral offered is owned by the borrower and is free of liens.

Covenants are a negotiated part of loan agreements. Warranties verify certain statements by the borrower at the date of execution of the loan agreement. Covenants carry forward the warranties and establish the borrower's ongoing obligation to maintain a certain status for the loan's duration. Covenants set minimum standards for a borrower's future conduct and performance and thereby accelerate the loan in the event of untoward developments. Violation of a covenant creates an *event of default* and gives the bank the right to "accelerate" the required repayment. We will have more to say about covenants in a later section of this chapter.

Informational Problems in Loan Contracts and the Importance of Loan Performance

Informational Problems

If there were no informational problems in loans, there might not be any profits for banks in lending. At one extreme, the costless availability of information obviates the need for banks and other financial intermediaries. At the other, costly customer-specific information provides an opportunity for banks to profitably process information and facilitate lending. In general, the less transparent is the credit information about a given borrower, the greater is the bank's ability to utilize its "uniqueness" and the higher is its profit potential. Thus, the paucity of good credit information in the public domain is a thing for banks to desire.[8] Since we have already discussed the informational problems addressed by banks (Chapters 2 and 3), we will merely review these here. The first problem is that the borrower is privately informed about its own credit risk. Unless the bank can elicit at least part of this information, market failure could result (recall the discussion of Akerlof in Chapter 1). We will see shortly that *credit analysis* helps the bank reduce its informational disadvantage vis-à-vis the borrower.

The other problem in lending is *moral hazard*. When the borrower takes a loan from the bank, it becomes an agent of the bank and is in a similar relationship with the bank as the shareholders of a firm are with bondholders. This agency problem is manifested in the borrower's desire to take on additional risk to the bank's detriment, as we saw in Chapter 1. Loan contracts are therefore designed to control the borrower's risk-taking propensity. To the extent that some preference for risk remains, the loan contract should also enable the bank to monitor the borrower and prevent actions that increase the risk of default. We will see how collateral, loan covenants, and other features of loan contracts can be structured to meet this important objective.

Figure 5.5 pictorially depicts the informational problems in loan contracts.

The Importance of Loan Performance

The bank's loan portfolio affects the financial health and viability of the bank. When bank stock prices decline, quite often most of the decline in bank stock prices is attributable to information releases about asset quality problems at banks.

8. Consider the following quote, *"Let us state a simple but often overlooked proposition: The health of a country's banking industry is inversely related to the speed and efficiency of information transfer,"* Sanford Rose, "Why Banks Make So Many Bad Loans," *American Banker*, June 19, 1990.

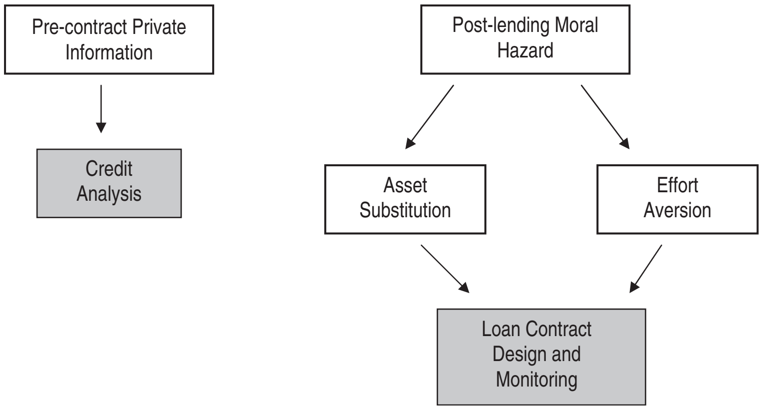


FIGURE 5.5 Information Problems in Loan Contracts

Loan losses can not only mean plunging stock prices, but can also spell trouble for top management at banks. It is well known that poorer corporate performance often precipitates a higher probability of CEO turnover.[9]

Loan Portfolio Diversification as a Risk Management Tool

The performance of a bank's loan portfolio often determines its financial performance. Although diversification is often a key to managing credit losses, many banks are constrained in their diversification efforts. For example, smaller banks often feel disadvantaged in their ability to diversify their credit risk by virtue of loan size limitations and geographic insularity. Moreover, these banks prefer limiting themselves to local markets because those are the markets they are most familiar with. This specialization-induced desire to stick to what is familiar leads to credit concentrations in banks, and this is typically reflected in the incidence of financial stress in periods of recession.

Despite regulatory attempts to encourage banks to diversify—by imposing limits on the maximum amount the bank can lend to a single borrower—the effects of lack of loan portfolio diversification can be clearly seen in the performance of banks in many regions. Banks in the United States have often displayed high performance correlation with banks that are similarly geographically situated. For example, when real estate values plunged in the 1980s in Texas, Oklahoma, and Louisiana, financial performance indicators for the banks in those states plunged as well.

It does appear, however, that the benefits of diversification have begun to more strongly influence banks' portfolio choices in the 1990s and post-2000, which coincides with the growing popularity of loan syndication, loan sales and securitization as diversification vehicles for banks. For example, although most United States community banks conduct much of their business in their own regions, there is recent evidence that these banks are able to withstand local economic downturns.[10] Moreover, in contrast to their relatively poor performance in the 1980s, small banks significantly

9. See, for example, Brickley (2003).

10. See John Hall and Timothy J. Yeager, *The Regional Economist*, "Does Relationship Banking Protect Small Banks From Economic Downturns?" *The Federal Reserve Bank of St Louis*, April 2002.

improved their performance in the 1990s. In fact, asset and deposit growth at small United States banks during the 1990s, when adjusted to account for the effects of mergers on measured growth, exceeded the growth at large banks.[11] In addition, small banks have also improved their profitability and survival rates. The FDIC reported that about 1,250 new community banks were established between 1992 and 2003, of which about 100 merged and about 1,100 remain independent, with only four having failed.

We now define terms that are routinely used in discussions of credit risk.

Interest Rate Spread: The difference between loan and deposit interest rates.

Provision for Loan Losses: A fraction of the loan principal earmarked by the bank as a buffer to absorb (expected) loan losses, and kept as part of the bank's capital.

Net Interest Spread After Provision: Interest rate spread after adjustment for taxes and subtraction of provision of loan losses.

Noninterest Income: Bank's income from activities other than lending, such as fees on cash management services, fees on contingent claims like loan commitments, letters of credit, and so on.

ROA: Bank's return on assets.

ROE: Bank's return on equity.

Nonperforming Loans/Reserves: Ratio of loans considered likely to default to the provision for loan losses.

Net Chargeoffs/Average Loans: Ratio of chargeoff of delinquent loans to the average loans extended by the bank.

Typically, interest rates are set such that interest rate spreads are higher for riskier loans. Banks also make higher provision for loan losses when the loans are riskier, and net chargeoffs/average loans also tend to be higher for such loans. Diversification can reduce the impact of losses in a particular loan class on the bank's overall net chargeoffs. Whether noninterest income, ROA and ROE are higher or lower for riskier loans depends on the degree of competition in that particular market and cannot be unambiguously stated *a priori*.

Despite the obvious gains from diversification, why are all banks not highly diversified? There are at least four reasons. First, there is the issue of limitations on the opportunity to diversify. Many banks feel "landlocked," constrained by geography to lend in limited markets. Second, lending opportunities typically arrive sequentially and unpredictably, so that forgoing a loan because of diversification concerns may be costly because a loan that offers better diversification potential may fail to materialize later. Third, banks are often constrained by regulations that mandate serving specific communities. For example, the *Community Reinvestment*

11. See William F. Bassett and Thomas F. Brady, "The Economic Performance of Small Banks, 1985–2000," *Federal Reserve Bulletin*, November 2001, pp. 719–728.

Act (CRA) requires a bank to lend to low-income borrowers in the community. This may interfere with diversification. Finally, cross-sectional reusability of information induces banks to specialize. For example, a bank that develops a special expertise in lending to auto parts manufacturers has a relative advantage in lending to this group, and it may wish to capitalize on this advantage by making such loans the focus of its loan portfolio. At best, therefore, banks tend to diversify within specialized areas of lending.

Credit Analysis: The Factors

Credit analysis examines factors that may lead to default in the repayment of a loan. The principal objective of credit analysis is to determine the ability and willingness of the borrower to repay the loan. The analysis looks at the borrower's past record (reputation) as well as its economic prospects. In most banks, this information is collected, analyzed, and stored by the credit department.

In analyzing a loan request, there are two important points to keep in mind. First, from an economic standpoint, assuming that the bank is the sole lender, it is the bank, not the borrower, that owns the asset financed with the loan. When the borrower takes a loan secured by the asset the loan is financing, it is merely purchasing a call option (as we saw in Chapter 1) from the bank. This option entitles the borrower to repurchase the asset from the bank should the value of the asset exceed the borrower's loan repayment obligation (the exercise price of the call option). The bank's loan granting decision and all of the actions it takes during the time the loan is outstanding should reflect this basic reality. Second, getting the borrower to repay the loan in today's legal environment is not always easy. Bankruptcy laws contain many provisions that protect borrowers, and these often make collection of debts potentially time-consuming and costly. Hence, one of the goals of credit analysis should be to uncover the likelihood of default as accurately as cost limitations will permit.

Traditional Factors Considered in Credit Analysis

Bank credit analysts have traditionally referred to the five Cs of credit analysis: capacity, character, capital, collateral, and conditions. Since "rules of thumb" are usually the distillate of accumulated experience, they should bear a relationship to theoretical prescriptions. We therefore, interpret each of these factors in terms of the underlying economics of bank lending. The discussion that follows is summarized in Figure 5.6.

(i) Capacity This refers to the borrower's legal and financial capacity to borrow. The first consideration in assessing a loan request is whether the person requesting the loan is legally capable of borrowing. For example, in the case of partnerships, it is important to know whether all the signing partners have the legal authority to borrow on behalf of the partnership. In the case of corporations, the bank should check the corporate charter and bylaws to determine who has the authority to borrow on the corporation's behalf.

Apart from legal considerations, capacity refers to the borrower's financial capability. Future cash flows are generally used to service the debt and therefore need to

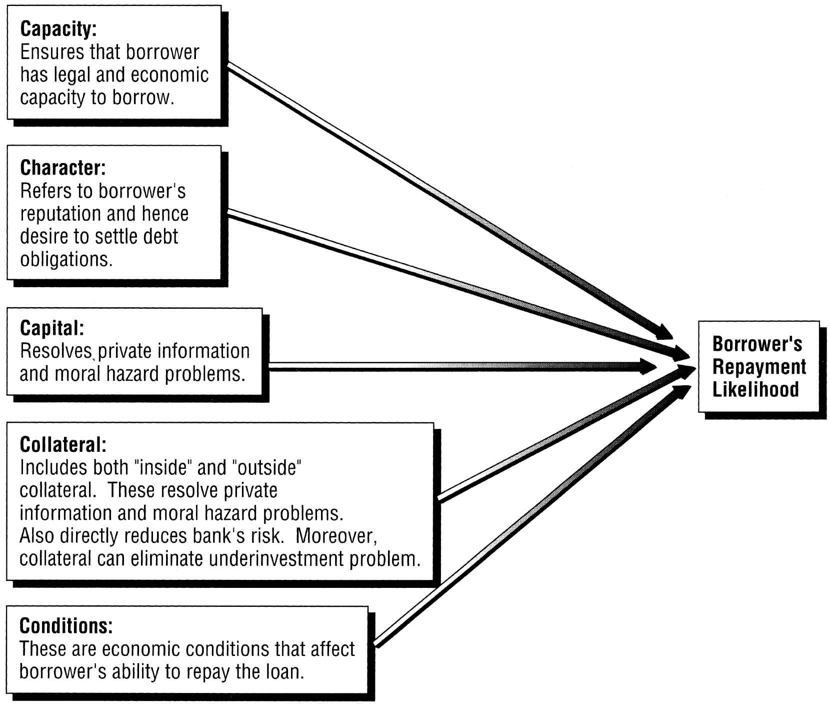


FIGURE 5.6 Pictorial Depiction of Factors Considered in Credit Analysis

be carefully estimated. Evaluating borrowers' future cash flows available to service the debt is a major part of any credit analysis. Sometimes, the bank may have to demand that the borrower subordinate the claims of others to ensure that the borrower has sufficient capacity to repay the bank. An example is a small firm that has borrowed significant amounts from its major shareholders.[12]

(ii) Character The concept of character embraces the borrower's ability to repay debts and the desire to settle all obligations within the terms of the contract. Judging character requires a careful examination of the borrower's past record in debt repayment and related behaviors. Including character in credit analysis makes sense because the better a borrower's credit reputation, the less incentive it has to default.[13] The reason is simple. Suppose a borrower knows that a single default will lead to denial of credit for a long time. The gain from defaulting is the amount the borrower does not repay the bank, but the gain from repayment is the net present value (NPV) of all the investment projects that might be financed with future bank loans; defaulting on this bank loan leads to a loss of that NPV. Clearly, this NPV increases as the interest rates on future loans decline. Further, the longer the borrower keeps repaying its loans, the better its credit reputation gets and the lower its future loan interest rates.[14] Hence, when the borrower acquires a good credit reputation, it perceives a

12. This is a case in which the bank may be successful in getting the borrower to subordinate the claims of earlier creditors. In general, this will be difficult as covenants on existing loans will generally prevent the borrower from taking such actions unilaterally.

13. This argument is formalized in Diamond (1989).

14. A better reputation leads to a lower interest rate because it becomes less likely that the borrower will default.

lower sequence of interest rates on its future loans than if it did not have that reputation. Consequently, the benefit of repaying the loan (or equivalently, the cost of defaulting) is greater for a borrower with a better reputation. To put it a little differently, the benefit of maintaining or building a reputation is greater the better the reputation is to start with. Hence, borrowers with better reputations (repayment records) tend to be better credit risks.

(iii) Capital How much equity capital (as a fraction of total assets) the borrower has invested in the firm is an important factor in the assessment of that firm's credit risk. There are two effects at work here. First, a higher amount of capital lessens the moral hazard problem. Second, the higher the capital, the better is the signal sent by the firm's owners about the confidence they have in the firm's future prospects. This helps to resolve the private information problem.

Example 5.1 Suppose you are a bank lending officer at the Midtown National Bank considering a loan request from Miller Manufacturing company for $1.05 million. The firm currently has $1 million in equity and its existing debt repayment obligation is $2 million. Assume that this equity is in the form of retained earnings invested in a noninterest-bearing account. The firm can invest the $1.05 million it will borrow from your bank in one of two projects (the bank cannot directly control which project the firm will invest in): A or B. Project A will yield a payoff of $2 million with probability 0.8 and $1 million with probability 0.2 at the end of the period. Project B will yield a payoff of $7 million with probability 0.2 and a payoff of zero with probability 0.8 at the end of the period. The firm's existing assets will yield a payoff of $3 million with probability 0.8 and a payoff of zero with probability 0.2 at the end of the period. The payoff on either project A or project B is statistically independent of the payoff on the firm's existing assets. These payoff distributions are common knowledge. For simplicity, there is no discounting and the bank loan you will make is subordinated to the firm's previous debt. Examine how Miller Manufacturing's behavior and the terms of lending change depending on whether or not it has the $1 million equity mentioned earlier.

Solution We solve this problem in four steps. First, we will assume that Miller Manufacturing has $1 million in equity. Then we will analyze the firm's expected profit from choosing project A, *assuming* that the bank prices the loan believing that project A will be chosen. Second, continuing to assume that Miller has $1 million in equity, we will analyze the borrower's expected profit from choosing project B assuming that the bank prices the loan believing that project A will be chosen. These two steps are needed to determine the appropriate Nash equilibrium in this problem, that is, a situation in which the bank prices the loan believing that Miller will choose project i (where i is either A or B) and Miller indeed chooses project i. With $1 million in equity, the Nash equilibrium involves the bank believing that project A will be chosen. Note the key role of the informational assumption that the bank cannot observe the borrower's choice of project. The third step is to assume that Miller has no equity capital and repeat Step 1. Finally, we repeat Step 2 with the assumption that Miller has no equity capital.

(*Continued*)

Step 1 Suppose first that the firm has the $1 million in equity mentioned earlier. Let's say that you, as the lending officer, assume that Miller will choose project A. Then, the *sum* of the cash flows from the project and Miller's existing assets has the following probability distribution.

TABLE 5.1 Probability Distribution of Total Cash Flows From Project A, Miller's Existing Assets and Equity

Total Cash Flow From Project A and Existing Assets Millions of $	Total Cash Flow With Retained Earnings Added in Millions of $	Probability
5	6	0.64
4	5	0.16
2	3	0.16
1	2	0.04

Since the repayment obligation on the senior debt is $2 million, the cash flow available to service the bank loan has the following probability distribution.

TABLE 5.2 Probability Distribution of Cash Flow Available to Service Bank Loan

Cash Flow Available Millions of $	Probability
4	0.64
3	0.16
1	0.16
0	0.04

You want to price this loan competitively because Miller has also been talking to your crosstown rival. At the same time, you do not want to lose money on this deal. From Table 5.2 you figure out that if the available cash flow is either $4 million or $3 million, Miller can fully repay the bank loan, whereas if the available cash flow is $1 million, then that is all your bank can collect. Thus, if you set the repayment obligation on your bank loan at $P million, your expected collection will be

$$(0.64 + 0.16)P + (0.16)1 + (0.04)0 = 0.8P + 0.16.$$

Since we've set the discount rate at zero, this expected payoff must equal the initial loan for your bank to just break even (the farthest you can go in competing for this borrower). That is,

$$1.05 = 0.8P + 0.16,$$

which means P = $1.1125 million, implying a loan interest rate of approximately 5.95 percent. The probability distribution of cash flows to Miller's shareholders is given in the table below.

TABLE 5.3 Probability Distribution of Net (Pretax) Cash Flow Accruing to Shareholders of Miller Manufacturing

Cash Flow Available Millions of $	Probability
2.8875	0.64
1.8875	0.16
0	0.16
0	0.04

Thus, the expected value of equity if Miller invests in project A is $0.64(2.8875) + 0.16(1.8875) = \$2.15$ million.

Step 2 Now suppose that Miller were to consider investing in project B after receiving a loan priced by you under the assumption that project A would be chosen. This is the standard moral hazard problem in bank lending that we discussed earlier, since project B is riskier for you as the lender. Then, proceeding in the same way that we did for project A, we see that the probability distribution of the cash flows accruing to the firm's shareholders is as follows: $7.8875 million with probability 0.16, $4.8875 million with probability 0.04, $0.8875 million with probability 0.64, and zero with probability 0.16. Thus, the expected value of equity if the firm invests in project B is $0.16(7.8875) + 0.04(4.8875) + 0.64(0.8875) + 0.16(0) = \$2.0255$ million.

This means that Miller's shareholders prefer to invest in project A (assuming you price your loan as if project A will be selected) and you are safe in your assumption that project A will be chosen. It is, therefore, unnecessary to check what would happen if the bank were to assume that Miller will choose B. This is because there are two possibilities. Either Miller will choose A, so that it is not a Nash equilibrium for the bank to assume B will be chosen, or there is a Nash equilibrium in which Miller chooses B. But this Nash equilibrium is dominated by the one in which Miller chooses A in the sense that Miller is better off in the latter and the bank is indifferent. Thus, if your bank is to be competitive, you had better price the loan assuming that A will be chosen, since the loan price is lower in that case.

Step 3 Now we will see what would happen if Miller had no equity capital. In this case, if Miller selects project A, it has the following distribution for its total cash flow.

TABLE 5.4 Probability Distribution of Total Cash Flows from Project A and Miller's Existing Assets

Total Cash Flow Millions of $	Probability
5	0.64
4	0.16
2	0.16
1	0.04

Since the repayment obligation on senior debt is $2 million, you calculate that to service the bank loan Miller will have $3 million with probability 0.64, $2 million with probability 0.16, and nothing with probability 0.2. Following the same logic as in the case with

(*Continued*)

$1 million in retained earnings, you now calculate that to permit the bank to just break even you must ask for a repayment obligation of $1.3125 million (you're assuming that Miller will invest in project A). With this, the net cash flow accruing to Miller's shareholders is $1.6875 million with probability 0.64, $0.6875 million with probability 0.16, and zero with probability 0.2. The expected value of equity is $1.19 million.

Step 4 After receiving such a loan, if Miller were to decide to opt for project B instead, we can follow the same steps as before to compute the expected value of equity as $1.2175 million. Thus, Miller will choose the riskier project B, and your assumption that it will select project A is incorrect. Indeed, if you were to (correctly) assume that project B will be chosen and price the loan accordingly, Miller's incentive to choose project B would be unaltered. This means that if Miller does not have sufficient equity capital, it may opt for riskier investments than it would if it had equity capital. Since you will anticipate this as a banker, you price the loan accordingly (that is, charge an appropriately higher interest rate on the loan). It is straightforward to verify that in this example Miller is better off retaining earnings in order to convince the lender that it will choose the safer project.

Capital helps to resolve moral hazard by imposing a greater loss on the borrower for poor project outcomes. This is because capital acts as the "first line of defense" against project losses and provides a cushion of protection for the lender. Without equity capital, the borrower knows that it has a valuable call option—if the project does poorly, the lender sustains the loss (the worst the borrower can do is to get nothing), whereas if the project does well, the lender gets only its contractual payment and the borrower earns a profit. With capital, the borrower's cost of pursuing risk is increased and the value of its call option is reduced. With sufficient equity capital, the lender can align the borrower's interest perfectly with its own. Interestingly, this means that the borrower is better off.[15]

The other function of capital is as an information communicator. The entrepreneur's own contribution of equity can signal the profitability of her project.[16] The standard argument relies on the entrepreneur being risk averse and is thus a little more complicated than an alternative line of reasoning that is developed in the example in the box below.[17]

Example 5.2 Suppose we have a firm that needs $150 to invest in a project that will yield a random payoff one period hence. The firm knows the probability distribution of the project's cash flow, but no one else does. All that others know is that the project can be type C or type D. If it is type C, then it will yield a cash flow of $300 with probability 0.8 and zero with probability of 0.2. If it is type D, the project will yield a

15. This is because of our assumption that the pricing of bank loans is competitive, so that the greater the equity capital possessed by the borrower, the better are its credit terms. Note that this provides an incentive for borrowers to accumulate equity capital.

16. See Leland and Pyle (1977).

17. This example is in the spirit of papers in the corporate finance literature that show a firm's choice of capital structure can signal its private information about its future prospects. See, for example, Ross (1977) and Shah and Thakor (1987).

cash flow of \$600 with probability 0.5 and zero with probability 0.5. For simplicity, suppose that interest and principal payments on debt are tax deductible and that the firm can raise equity capital (it currently has negligible equity capital on its books) from those who know the firm's cash flow distribution (for example, these may be managers who own stock). The firm currently has owners, but the book value of their equity is, for all practical purposes, zero. However, debt must be acquired in the form of a loan from a bank, which cannot tell whether the borrower has a type C or a type D project. The corporate tax rate applicable to the borrower is 30 percent. As a banker, how should you deal with such a borrower, assuming that the borrower is locked into either project C or project D and cannot choose its project?

Solution The key to resolving this informational asymmetry is to use capital as a signal. As a banker, the key is for you to recognize that the riskier borrower has a greater aversion to putting up equity capital because he has a greater likelihood of losing it. So, as a banker, you can offer the borrower two choices: (i) borrow the entire \$150 and repay P_D, or (ii) put up \$E in equity, borrow \$150 − E and repay P_c.

We solve this problem in three steps. First, we assume that the type-D borrower opts for choice (i), the type-C borrower opts for choice (ii), and the bank earns zero expected profit on each borrower. We then solve for P_D. We also solve for P_c, but it appears as a function of E. Step 2 involves solving for E. We do this by searching for the smallest value of E that ensures that the type-D borrower does not prefer its own contract (borrowing without putting up any equity) to that of the type-C borrower (putting up E in equity). Finally, the third step is to check that, with the value of E obtained from the previous step, the type-C borrower prefers his choice to that of the type-D borrower. Steps 2 and 3 therefore confirm the assumptions made in Step 1 about the project choices of borrowers.

Step 1 Now, if borrowers self-select so that only the type-D borrower takes (i) and only the type-C borrower takes (ii), then we can proceed as follows. Given that the bank must earn zero expected profit on each contract, and the repayment probability of the type-D borrower is 0.5, P_D must equal the expected value of the bank's repayment by the high-risk borrower, that is,

$$P_D \times 0.5 = 150$$
$$\text{or } P_D = \$300, \text{ an interest rate of 100 percent.}$$

Next, if only the low-risk borrower takes (ii), P_C must satisfy

$$0.8 \times P_C = 150 - E$$
$$\text{or } P_C = \frac{150 - E}{0.8}.$$

Step 2 We now solve for E. Note that E must ensure that the type-D borrower does not prefer the type-C borrower's contract to his own. Although there are many values of E for which this is true, there is only one value of E for which this is true and the value of the debt tax shield for the type-C borrower is maximized. This is the value of

(Continued)

E that is the smallest value such that the type-D borrower does not strictly prefer the type-C borrower's contract. That is, the NPV to the type-D borrower from misrepresenting [and choosing (ii)] is exactly equal to his NPV from telling the truth [and choosing (i)]. The type-D borrower's NPV from choosing (i) is

$$(600 - 300) \times 0.5 \times 0.7 = \$105$$

where 0.7 is one minus the tax rate. The type-D borrower's NPV from choosing (ii) is

$$\left(600 - \frac{150 - E}{0.8}\right) \times 0.5 \times 0.7 - E.$$

Equating the above NPV to \$105 yields E = \$70. Thus, the repayment obligation for the type-C borrower is $\frac{150 - 70}{0.8} = \$100$, or an interest rate of 25 percent.

Step 3 You can check that the type-C borrower will strictly prefer his contract to that of the type-D borrower. His NPV from (i) is

$$(300 - 300) \times 0.8 \times 0.7 = 0,$$

and his NPV from (ii) is

$$(300 - 100) \times 0.8 \times 0.7 - 70 = \$42.$$

Thus, the bank can offer two choices:

(i) Borrow the entire \$150 and repay \$300.
(ii) Put up \$70 in equity, borrow \$80, and repay \$100.

The key here is that the bank prices each loan based on the assumption that the borrower taking a particular loan has a particular project. If the borrower does in fact have that project, then the bank earns zero expected profit. The idea is for the bank to design the loan in such a way that *incentive compatibility* is assured. In other words, no borrower has an incentive to deviate from the loan contract "intended" for it by the bank. Incentive compatibility should obtain in a *Nash equilibrium*; the bank's assumptions about the association between the borrower's project and its loan contract choice must be correct in equilibrium.

In this example, capital serves as a signal of project quality. The borrower with the less risky type-C project signals its lower risk by funding two-thirds of the required investment with equity capital. For this, it is rewarded with a lower interest rate. Despite the obvious attractiveness of this lower interest rate, the high-risk borrower is unwilling to put up the equity necessary to be granted that rate. The intuition is as follows. Due to the tax deductibility of loan interest payments, the borrower desires as large a loan as possible, *regardless* of its project characteristics. The borrower also dislikes paying interest, regardless of its project characteristics. However, a higher interest rate is less onerous when the borrower has a risky project because the likelihood of actually repaying the loan with interest is lower. To such a borrower then, the inducement of a lower interest rate in exchange for a higher capital requirement is less attractive than it is to a borrower with the safer project. It is the fact that the borrower's preferences over different capital requirement-interest rate combinations *depend* on its project characteristics that permit the bank to craft a self-selection mechanism that elicits the desired information.

It follows then that, all else remaining the same, the bank should charge an interest rate that is *inversely* related to the borrower's equity to total assets ratio. Less capitalized borrowers are more risky, not just because of the *direct effect* of capital in serving as a "first line of defense," but also because of its *indirect effect* in reducing the borrower's appetite for risk. In our examples, we imposed a zero-profit condition on the bank as a reflection of perfect competition in banking markets. This is an extreme representation of competition. In reality, banks earn profits, especially on borrowers about whom they possess credit information that is not publicly available. To the extent that banks charge higher interest rates to borrowers with lower equity positions, they may also be able to earn greater profit margins on these borrowers.[18] This can make the prospect of lending to highly leveraged (low equity) borrowers enticing for the bank, despite the higher risk involved. Indeed, such an incentive arises from the basic function of credit information production performed by banks (Chapter 3).

Banks can add highly leveraged loans to their portfolios by lending to companies that use the funds for leveraged buyouts, acquisitions, and recapitalizations. As our earlier discussions indicate, the yields on these highly leveraged transactions (HLTs) are higher than on other commercial loans. Since these higher yields compensate the bank for higher risks, higher *expected* profits for the bank are not necessarily implied. However, in many cases these borrowers also have few alternative sources of credit, so that banks can extract higher risk-adjusted profits from these borrowers. In addition, banks usually receive fees that vary from one to two percent of the principal amount committed.[19] HLT loans, however, are significantly more risky than average, and involve the moral hazards discussed earlier in this section.[20] This may be one reason why there has been a recent growth in the popularity of *reverse leveraged buyouts*, whereby firms reduce their debt/equity ratios by issuing equity to retire debt acquired during leveraged buyouts (LBOs). This would reduce moral hazard and benefit the firm.

(iv) Collateral Most commercial and consumer lending is secured with collateral.[21] Once a loan is secured by a specific asset that serves as collateral, the lender has first

18. This may also be because borrowers with lower equity capital levels may be less well known and have access to fewer credit sources, so that banks can earn higher quasi-monopoly rents by producing private information about them.

19. Usually, these loans are made under loan commitments, so that the fees are commitment fees.

20. An HLT loan may not only impose a higher expected loan loss for the bank but may also involve higher loan loss volatility (see Chapter 6).

21. For example, based on the Federal Reserve's Survey of Terms of Bank Lending, Boot, Thakor, and Udell (1991) report that, as of May 1988, 69.1 percent of bank loans were secured. See also Jimenez, Salas and Suarez (forthcoming).

claim to that asset in the event of default. There are two types of collateral: "inside" and "outside." Inside collateral consists of assets owned by the firm to which the loan is extended. Examples are accounts receivables, equipment, machinery, real estate, and inventory. Even if the bank extends an unsecured loan, it would have a claim, but not necessarily first claim, against these assets. As a general creditor, however, the value of the bank's claim would be ill-defined since, in the event of bankruptcy, the bank might be one among many unsecured creditors at the mercy of the bankruptcy court. On the other hand, if one of these assets is pledged as (inside) collateral, the bank would become the primary claimant to that asset.

Outside collateral consists of assets that the bank would never have a claim to unless they were specifically designated as collateral. A good example would be personal assets of the owner of the borrowing corporation or limited partnership.

Using collateral is not costless, however. Since the borrower may undertake actions that undermine the value of the collateral to the bank, ongoing monitoring of the collateral is required. Such monitoring costs are absorbed, at least in part, by the bank. Moreover, when collateral is transferred to the bank upon default, there are liquidation costs. These include the legal costs of ownership transfer as well as the bank's costs of initially carrying and then selling off the collateral.[22] From the borrower's standpoint, use of collateral makes subsequent borrowing more expensive since fewer assets are available to general creditors on that borrowing. Despite these costs, why is collateral so widely used?

There are at least three reasons for the popularity of secured lending. We discuss each now.

(a) Risk Reduction: An obvious reason to secure a loan is that it provides the lender greater protection against loss in the event of default. The bankruptcy code in the United States includes what is known as an "automatic stay," which freezes collection actions by creditors during bankruptcy proceedings. The idea is to provide the debtor with breathing room to put its house in order. The stay takes effect immediately upon the filing of a bankruptcy petition. However, the stay can be modified in favor of a creditor if there is "cause," including insufficient protection of the *secured* creditor's interest in that component of the debtor's property that serves as collateral. For example, suppose a bank has loaned \$10 million to a firm that has just filed for reorganization under Chapter 11 of the bankruptcy code. Suppose that specific assets of the firm, currently worth \$4 million, have been encumbered as inside collateral. Now, if these assets were to depreciate in value at the rate of \$3,000 per month, for instance, the bankruptcy court might require the firm to set aside that amount each month to adequately protect the bank's claim. Thus, securing a loan reduces the creditor's risk in the event of bankruptcy.

(b) Signaling Instrument: Collateral can also convey valuable information to the bank. Although possible with inside collateral, the intuition comes through most clearly if one thinks of securing property as outside collateral. The logic is similar to that used in explaining the signaling role of equity capital. Within a class of borrowers that look equally risky to the bank even after all credit analysis is done, a borrower's willingness to offer collateral will be inversely related to its

22. By regulation, banks are required to liquidate such holdings within a certain time period after acquisition, unless the collateral is a permitted bank asset holding.

default risk on the loan.[23] The way the bank can induce a borrower to reveal its otherwise hidden risk is as follows. Suppose there are two indistinguishable borrowers, A and B. However, the bank suspects one may be riskier than the other, although it does not know which. The bank offers each borrower a choice of one contract from a pair consisting of a secured loan with an associated interest rate and an unsecured loan with a higher interest rate. Now suppose A is less risky than B. Then, A will prefer the secured loan for two reasons. First, its lower risk means that the likelihood of repaying interest is higher; hence, a lower interest rate is more appealing. Second, its lower risk means that the chance of defaulting and losing collateral to the bank is lower; hence, offering collateral is less onerous. By symmetric logic, we can see that B will prefer the unsecured loan. Getting A and B to sort themselves out like this requires, of course, that the two loan contracts offered are incentive compatible. The example in the box below shows how this can be done.

Example 5.3 Suppose that A's assets will be worth \$100 for sure at the end of the period. The value of B's end-of-period assets will be \$200 with probability 0.5 and zero with probability 0.5. The project (A or B) requires an investment of \$30 up front and the entire amount is borrowed from the bank. The bank is unable to distinguish between A and B. Assume that the single-period riskless interest rate is 10 percent and everybody is risk neutral. Assume that collateral worth \$1 to the borrower is worth only 90 cents to the bank. The difference of 10 cents on the dollar can be viewed as the bank's cost of taking possession of the collateral. These repossession costs have two sources. First, assets acquired from a delinquent borrower are often worth less piecemeal to the bank than they are to the borrowers as components of a productive whole. Thus, the mere act of liquidating collateral by removing it from the other assets of the firm is costly. Second, transferring control of assets from the borrower to the bank involves legal and other administrative costs. These costs are an important reason why so many bankers see the value of collateral largely in terms of its incentive effects. The problem is to determine how the bank can design a *pair* of loan contracts such that each borrower will be induced to truthfully reveal its privately known risk.

Solution Following the intuition discussed earlier, we will need to offer borrowers two contracts: a secured loan and an unsecured loan. These contracts should be designed so that A, the safe borrower, chooses the secured loan and B, the risky borrower, chooses the unsecured loan. We solve this problem in three steps. In the first step, we solve for the interest rate on the secured loan for the bank to break even. Second, we solve for the interest rate on the unsecured loan. In the third step, we solve for the amount of collateral on the secured loan that will deter the risky borrower from preferring the secured to the unsecured loan.

(*Continued*)

23. See Bester (1985), Besanko and Thakor (1987a, 1987b), and Chan and Thakor (1987) for theoretical models that demonstrate this. Empirical evidence on the signaling role of collateral is provided by Jimenez, Salas and Suarez (forthcoming).

Step 1 Since A will surely repay the loan, the interest rate on the secured loan, r_u, that allows the bank to just break even is the single-period riskless rate of 10 percent.

Step 2 On the other hand, the interest rate on the unsecured loan, r_u, should be set to satisfy the following zero profit condition for the bank

$$[0.5 \times (1 + r_u) \times 30]/[1.10] = 30 \qquad [5.1]$$

The left-hand side of (5.1) is the discounted present value of the bank's payoff. The promised repayment is $\$30(1 + r_u)$, but there is only a 0.5 probability that the bank will be repaid. Since the bank is risk neutral, it discounts at the riskless interest rate of 10 percent. For the bank to exactly break even, the discounted present value of its expected payoff should exactly equal the initial loan. Note that our approach is consistent with the notion that the bank owns the project and it has sold the borrower a call option on the collateral at a fixed exercise price of $30 \times (1 + r_u)$. When the project value exceeds this exercise price, the borrower exercises the option to repurchase the project; this happens in the successful state. If the project fails, the borrower lets its option expire unexercised and the bank retains a worthless project. Solving (5.1) gives $1 + r_u = 2.2$. Hence, the repayment obligation on the unsecured loan is $2.2 \times 30 = \$66$.

Step 3 Now we solve for the amount of collateral that will deter B from mimicking A and opting for the secured loan. The amount of collateral, C, that makes B indifferent between the secured and unsecured loans is the solution to the following equation

$$0.5 \times (200 - 66) = 0.5 \times (200 - 33) - 0.5 \times C. \qquad [5.2]$$

In (5.2), the left-hand side is the expected value of the borrower's cash flow, net of repaying the bank, if it takes the unsecured loan. The right-hand side is the expected value of its net cash inflow if it chooses the secured loan. Note that the interest rate on the secured loan is 10 percent (since the bank assumes this loan will be taken by the safe borrower), so that the repayment obligation is $1.10 \times 30 = \$33$. There is a 0.5 probability that the borrower will default and lose its collateral to the bank.

Solving (5.2) yields $C = \$33$. Thus, if the bank demands a collateral whose value to the borrower is at least as great as \$33, only A will choose the secured loan with an interest rate of 10 percent. (Note that A's net expected cash flow with the secured loan is $\$100 - \$33 = \$67$, whereas with the unsecured loan it is $\$100 - \$66 = \$34$). B will choose the unsecured loan with an interest rate of 120 percent. The bank can thus sort its borrowers according to risk. The outcome is a Nash equilibrium; the bank's beliefs about which borrower chooses which loan is confirmed by their behaviors.

You must have noticed that the bank's collateral repossession cost had no bearing on the outcome. The reason is that the secured loan to A is *riskless*, so that A would never surrender collateral to the bank. Since the Nash equilibrium separates perfectly—each borrower revealing its type in equilibrium—and involves B choosing the unsecured loan, the bank never actually takes possession of collateral in this example. In reality, of course, few loans are riskless. With default risk in lending to A, then the bank's repossession cost would have entered the outcome since it would have affected the interest rate on the secured loan.

(c) **Moral Hazard:** Using collateral can help resolve a variety of moral hazard problems. The three we will discuss here are: asset substitution, underinvestment, and inadequate effort supply.

Asset Substitution: Because of the option nature of the bank loan, the borrower has an incentive to choose a riskier project after obtaining the loan. In a manner similar to capital, collateral can deter such risk-taking. For present purposes, think of security offered as outside collateral. Consider the following example.

Example 5.4 Suppose Brown Bakery needs a \$100 loan to finance a project that will pay off next period. Brown can choose between two projects: S (safe) and R (risky). The bank knows this but is unable to directly control the borrower's choice of project. S will yield a payoff of \$300 with probability 0.9 and nothing with probability 0.1, and R will yield a payoff of \$400 with probability 0.6 and nothing with probability 0.4. Everybody is risk neutral and the riskless rate is 10 percent. How should the bank design its loan contract so that Brown will choose the safer project? Assume once again that collateral worth \$1 to Brown is worth 90 cents to the bank.

Solution The idea is for the bank to make it in Brown's best interest to choose S. This is achieved by demanding that Brown put up sufficient collateral. Since collateral is surrendered to the bank upon default, it makes project failure costly to the borrower. Consequently, the borrower will wish to minimize the likelihood of failure by choosing S. The key assumption here is that the bank cannot *directly* control Brown's project choice. We proceed in four steps. First, we will assume that the bank offers Brown an unsecured loan, assuming that S will be chosen. We will show that this cannot be a Nash equilibrium because Brown will choose R. Second, we will let the bank assume that R will be chosen and compute the interest rate on the unsecured loan. It turns out this is a Nash equilibrium in that Brown chooses R when faced with such an unsecured loan. Third, we ask whether another Nash equilibrium is possible, say with a secured loan. We solve for the level of collateral that ensures that Brown does not (strictly) prefer R to S. We do this by equating Brown Bakery's expected profits from R and S, given a secured loan contract will indeed be acceptable to Brown Bakery and the bank. Finally, we verify that it is a Nash equilibrium for Brown to choose S.

Step 1 First suppose the bank offers Brown an unsecured loan at an interest rate r_u. If the bank assumes that Brown will choose S, then the interest rate, r_u^S, at which the bank just breaks even, is given by

$$[0.9 \times (1 + r_u^S \times 100)]/[1.10] = 100. \qquad [5.3]$$

(*Continued*)

Solving (5.3) yields $r_u^S = 22.22$ percent. Can this be a Nash equilibrium in the sense that Brown does indeed choose S? To answer this question, let us compute Brown's net expected payoffs under R and S. If Brown chooses S, its net expected payoff is

$$0.9[300 - (1.22 \times 100)] = \$160.20.$$

If it chooses R, its net expected payoff is

$$0.6(400 - 122) = \$166.8.$$

Hence, offering Brown an unsecured loan with an interest rate of 22 percent cannot be a Nash equilibrium since Brown will choose R instead of S, and the bank will make an expected loss on the loan since it assumed S would be chosen.

Step 2 Now suppose the bank assumes that R will be chosen. Then the interest rate, r_u^R, at which the bank just breaks even, is given by

$$0.6 \times (1 + r_u^R) \times 100]/[1.1] = 100. \qquad [5.4]$$

Solving (5.4) yields $r_u^R = 83.33$ percent. Now, confronted with this interest rate, if Brown chooses S, its net expected payoff is

$$0.9(300 - 186.33) = \$105.$$

If it chooses R, its net expected payoff is

$$0.6(400 - 183.33) = \$130.$$

So, Brown chooses R and this is a Nash equilibrium since the bank's belief is consistent with the borrower's behavior.

Step 3 But can we do better with *another* Nash equilibrium? Whenever we ask this question, it is natural to wonder who we are doing better for. Since the bank is assumed to earn zero expected profits in all scenarios, why should the bank care? The answer lies in competition. Recall that the zero expected profit condition is an analytical convenience. In practice we would expect the bank to earn at least a small profit. Remember too that this profit is in excess of the normal return on equity capital. Now, if the bank can design a contract that increases the borrower's expected profit without reducing the bank's, it can lure away this borrower from its competitors and build its "book" of business. Hence, competing banks should strive to give the borrower the best possible deal.

Suppose now that the bank offers Brown a secured loan instead. What you want to do as a banker is to figure out how much collateral to ask for in order to ensure that R will not be chosen. The level of collateral that leaves Brown indifferent between S and R satisfies the following equation.

$$\begin{aligned} &0.9[300 - (1 + r_s) \times 100] - 0.1 \times C \\ &\quad = 0.6[400 - (1 + r_s) \times 100] - 0.4C. \end{aligned} \qquad [5.5]$$

where r_s is the interest rate on the secured loan. We should first determine r_s. If the bank is successful in inducing Brown to choose S, then it should set r_s as follows to satisfy its zero profit condition

$$[0.9 \times (1 + r_s) \times 100 + 0.1 \times 0.9 \times C]/[1.1] = 100. \qquad [5.6]$$

In (5.6), note that we have used the fact that a dollar of collateral is worth only 90 cents to the bank. Solving (5.6) yields

$$1 + r_s = (110 - 0.09C)/90. \qquad [5.7]$$

Substituting (5.7) in (5.5) and solving for C yields C = \$20,202. To avoid rounding off problems, suppose we take C = \$20.21. Then substituting this in (5.7) gives us $1 + r_s = (110 - 1.8189)/90 = 1.2020$ or say $r_s = 20.21$ percent to make sure that rounding off does not leave the bank with negative expected profit.

Step 4 Now Brown's net expected payoff from choosing S is [from (5.5)] \$159.79 and from choosing R it is [again from (5.5)] approximately \$159.79. Hence, this is a Nash equilibrium in which Brown chooses S. Note that this equilibrium gives Brown a higher expected payoff than the previous Nash equilibrium (\$130).[1] Thus, if this borrower comes to you and says that your crosstown rival has offered an unsecured loan at 83.33 percent interest, you could effectively counter by offering a secured loan that requires \$20.21 of outside collateral and an interest rate of say 21 percent. With these terms, Brown Bakery will accept your loan and you will earn a profit.[2]

1. As noted in Chapter 1, there are often multiple Nash equilibria.
2. By this time, you may be wondering why a bank would ever make an *unsecured* loan. Note, however, that offering *both* secured and unsecured loans helps to resolve private information problems. Moreover, it is not *always* optimal to use outside collateral to resolve moral hazard. Indeed, in Example 5.4, if the payoff in the successful state for project R is \$500 instead of \$400, the best outcome is for the bank to offer an unsecured loan priced under the assumption that R will be chosen.

In this example, outside collateral was used since we assumed limited liability, that is, it would not be lost upon bankruptcy if it were not pledged. For somewhat different reasons, inside collateral can also deter asset substitution. By securing specific assets within the firm, creditors can ensure that these assets will not be replaced by those that increase the risk exposure of creditors. Since this reduction in asset substitution possibilities will be reflected in a better price for the firm's debt, the advantage of issuing secured debt accrues to the firm's shareholders.[24]

24. The argument that inside collateral can help in this way to resolve asset-substitution problems was made by Jackson and Kronman (1979) and Smith and Warner (1979).

Underinvestment: One manifestation of the divergence of interests between the borrower and the lender is in the borrower being unwilling to invest additional funds in a project even though doing so increases the total NPV of this project.[25] The intuition is simple. Suppose you own some real estate that was financed mainly with a bank loan; this real estate is currently worth \$1.5 million. You could spend an additional million dollars that would enhance the real estate's value by \$1.1 million. However, suppose that the present value of your repayment obligation to the bank is \$2 million. Then, although investing \$1 million yields an NPV of \$100,000 for the project as a whole, it is not a good idea for you, the owner/borrower. This is because you increase the present value of the cash flows accruing to you by $\$(1.5 + 1.1)$million $-$ \$2 million $=$ \$600,000, but it costs you \$1 million, that is, the investment has a *negative* NPV of \$400,000 to you (the borrower), but a positive NPV of \$100,000 to the borrower and lender considered jointly. The net effect is that the investment is passed up and firm value is sacrificed. This investment inefficiency arises from actions that are privately optimal for the borrowing firm's shareholders *ex post*. However, they pay a price for this *ex ante* since the lender anticipates such behavior and adjusts the terms of credit accordingly. How can we eliminate this form of moral hazard so that the *borrower* benefits *ex ante* through better credit terms?

One answer is to let the borrower *precommit* not to "underinvest" ex post. If the lender believes the borrower, the problem will have apparently been solved. However, such precommitment is *time inconsistent*. The lender knows that the borrower has every reason to break this promise when the opportunity presents itself. So it would be foolish for the lender to believe such a promise. Of course, loan covenants can be employed, with the lender monitoring compliance. However, as a practical matter, it is difficult to see how loan covenants could force a borrower to invest when it is disinclined to do so. This is because the lender typically does not "see" these investment opportunities unless the borrower decides to exploit them. Covenants are effective in *prohibiting* actions, but rarely succeed in forcing unobservable initiatives.

Secured debt can resolve this underinvestment problem.[26] The idea is as follows. Suppose that the firm needs additional financing to purchase an asset, and it can purchase this asset for less than its market value. Thus, the purchase is a positive NPV investment. Also suppose that the firm currently has risky unsecured debt outstanding and would not, without further incentive, purchase this asset because it would enhance the present value accruing to the firm's shareholders by less than the purchase price of the asset.

To solve this problem, suppose the firm issues new debt secured by the asset in question. Then, due to the "absolute priority" rule, the secured creditors have first claim to the asset in the event of bankruptcy, and the borrowing firm has essentially diverted (at least part of) the cash flows attributable to this asset to the new secured creditors and away from the old unsecured creditors. Since the new (secured) creditors pay a fair market value for the debt issued by the firm, the gains associated with diverting payoffs of the newly purchased asset away from the old (unsecured) creditors accrue to the borrowing firm's shareholders and increase their incentive to undertake the investment. The example in the box below illustrates how this works.

25. This underinvestment problem is discussed by Myers (1977).
26. This point was made by Stulz and Johnson (1985).

Example 5.5 Consider a firm, Johnson Supplies, that can invest \$100 at the start of the period ($t = 0$) in a project that will pay off at the end of the period ($t = 1$) \$400 if successful (state S_1) and zero if unsuccessful (state S_2). State S_1 occurs with probability 0.7. The initial \$100 financing comes from unsecured debt issued at $t = 0$. Before the end of the period, but *after* the initial financing is raised, the firm will have an opportunity to purchase an asset (call it A) for \$100. This asset will surely be worth \$120 at $t = 1$. Assume that Johnson cannot be forced to purchase this asset.[1] Compute Johnson's optimal financing strategy. Assume that everybody is risk neutral and that the riskless interest rate is 10 percent.

Solution We solve this problem in six steps. First, we assume that only unsecured debt can be offered and that the date-0 unsecured creditors will assume that Johnson will purchase A when available. We then compute the interest rate on the \$100 of (new) unsecured debt raised (*after* the initial financing) to purchase A. Second, we check if this can be a Nash equilibrium. We find that it is not, in that Johnson will *not* purchase A when burdened with the original unsecured debt. Third, we check if it is a Nash equilibrium for Johnson not to purchase A. That is, if the original creditors price their debt assuming that Johnson will not purchase A, will Johnson indeed not purchase A (since Johnson does not purchase A, we need not worry about the old creditors)? We find that this is a Nash equilibrium. Fourth, we introduce secured debt and compute the interest rates on the old unsecured and the new secured debt when all creditors assume that Johnson will purchase A when available. Fifth, we check if this is a Nash equilibrium. We find that it is a Nash equilibrium in that Johnson does purchase A and also wishes to issue secured debt to purchase A. Finally, in step 6 we conclude by indicating that the NPV to Johnson's shareholders is higher in the secured-debt Nash equilibrium than in the unsecured-debt Nash equilibrium when Johnson does not purchase A.

Step 1 First suppose that issuing secured debt is impossible. Thus, the \$100 financing required to purchase A in the future will have to be raised with either equity or unsecured debt. Since the basic argument follows in either case, let us assume that unsecured debt will be employed. As a start, suppose the unsecured creditors at $t = 0$ (call them C_{old}) assume that Johnson will purchase A when available. Use C_{new} to label the (new) unsecured creditors who provide the \$100 to buy A. Thus, at $t = 1$, the value of the firm will be \$520 (in state S_1) with probability 0.7 and \$120 (in state S_2) with probability 0.3. Assuming that all unsecured creditors have equal priority, C_{old} will be repaid in full in state S_1 and will receive \$60 in state S_2. The payoffs to C_{new} are identical. Hence, the loan interest rates on the credits provided by C_{old} and C_{new} will also be identical. Let r_a represent this interest rate. Then, if creditors provide fairly priced debt (that is, each creditor earns zero expected profit), r_a is obtained as a solution to the following equation

$$100 = [(1 + r_a) \times 100 \times 0.7 + 60 \times 0.3]/[1.1]. \qquad [5.8]$$

The left-hand side of (5.8) is the amount of debt financing. The right-hand side is the expected payoff to either C_{old} or C_{new}, discounted at the riskless rate of 10 percent. Solving (5.8) yields $r_a = 31.43$ percent. Thus, at $t = 1$ Johnson is obliged to repay \$131.43 to C_{old} and the same amount to C_{new}.

(*Continued*)

Step 2 The first question is: Can this be a Nash equilibrium? To answer this, we must find out whether C_{old}'s assumption that Johnson will purchase A is indeed correct. Now, if Johnson purchases A, the NPV accruing to its shareholders is

$$\frac{0.7 \times (520 - 262.86)}{1.1} = \$163.63.$$

Note that Johnson's shareholders receive a positive payoff only in state S_1, and this payoff is \$520(\$400 + \$120) minus two times \$131.43, where \$131.43 is what Johnson owes each group of unsecured creditors. If, on the other hand, Johnson does not purchase A, then the NPV accruing to its shareholders is

$$\frac{0.7 \times (400 - 131.43)}{1.1} = \$170.91.$$

Thus, Johnson will forgo the opportunity to purchase A even though its total NPV (\$120 – \$100/1.1 = \$18.18) to Johnson is positive. This means that it *cannot* be a Nash equilibrium for C_{old} to assume that Johnson will purchase A.

Step 3 So now suppose C_{old} assumes that Johnson will *not* purchase A. Then, the loan interest rate, r_b, is a solution to

$$[0.7 \times (1 + r_b) \times 100]/[1.1] = 100 \qquad [5.9]$$

Solving (5.9) yields $r_b = 57.143$ percent. It is simple to verify that, faced with this loan interest rate, Johnson will indeed choose *not* to purchase A. Thus, this is a Nash equilibrium, under the assumption that secured debt is impossible. The NPV accruing to Johnson's shareholders in this Nash equilibrium is given by

$$\frac{0.7 \times (400 - 157.143)}{1.1} = \$154.5.$$

Step 4 Imagine now that Johnson is free to finance A with secured debt. If Johnson chooses to do this, then the (secured) claim of C_{new} will be riskless since the minimum firm value (that prevails in state S_2) is \$120 (the value of A at $t = 1$), and C_{new} have first claim to this asset. Since the riskless rate is 10 percent, Johnson's repayment obligation on riskless debt will be \$110, and this can be covered from the value of this firm in state S_2. Now suppose C_{old} assumes that Johnson *will* purchase A when available. The loan interest rate, r_c, that C_{old} charges will then be a solution to

$$[0.7 \times (1 + r_c) \times 100 + 0.3 \times 10]/[1.1] = 100, \qquad [5.10]$$

where we recognize that C_{old} will be paid only \$10 in state S_2 since C_{old}'s claim is subordinated to that of C_{new}. Solving (5.10) gives us $r_c = 52.86$ percent. Johnson's total repayment obligation, therefore, is \$152.86 + \$110 = \$262.86.

Step 5 Is this a Nash equilibrium? Again, we consider Johnson's incentive to purchase A. If it purchases A, the NPV accruing to its shareholders is

$$\frac{0.7 \times (520 - 262.86)}{1.1} = \$163.63.$$

and if it does not purchase A, the NPV accruing to shareholders is

$$\frac{0.7 \times (400 - 152.86)}{1.1} = \$157.3.$$

Hence, Johnson will indeed purchase A (when C_{old} prices the loan assuming A will be purchased) and the conjecture of C_{old} about the firm's incentive to purchase A is supported by its behavior. To complete our verification that this is a Nash equilibrium, we must also make sure that Johnson will indeed wish to issue secured debt to purchase A. To check this, let us hold the fixed price of the loan given by C_{old}, so that the firm must repay \$152.86. If Johnson issues unsecured debt to purchase A, then C_{new} will ask for a loan interest rate of 31.43 percent [since they solve (5.8) to determine this loan interest rate], so that the NPV accruing to Johnson's shareholders is

$$\frac{0.7 \times [520 - (152.86 + 131.43)]}{1.1} = \$150.$$

Step 6 Thus, Johnson will indeed choose to finance A with secured debt. Moreover, the NPV to Johnson's shareholders in this Nash equilibrium (\$163.63) exceeds that in the previous Nash equilibrium when it could only finance the purchase of A with unsecured debt (\$154.5). Hence, it will *not* be in the interest of Johnson Supplies to precommit to never issue *secured* debt in the future through restrictive covenants written into its loan contract with C_{old}.

1. A simple way to ensure this is to assume that the opportunity to purchase the asset will arrive with some probability *less than one* and that creditors are unable to observe whether this opportunity has indeed arrived. This will not change the basic argument, but will complicate the numerical example a bit.

Apart from illustrating how secured debt can resolve the underinvestment problem, this example brings up an interesting point related to the design of covenants in loan contracts. It is sometimes believed that creditors wish to protect themselves against future expropriation by including loan covenants that prohibit the firm from issuing future debt that has a higher seniority claim against any subset of the firm's assets. When all is said and done, however, in a competitive market it is the *borrower* who decides what covenants to accept, since the lender can presumably adjust the price of the loan (to at least break even) depending on the covenants that the borrower is willing to accept. What our example shows is that it *may* be optimal for the borrower to leave itself the flexibility to avail of secured borrowing in the future in which the newly purchased assets are used as collateral, so that new creditors have the most senior claim to the assets.[27] This not only makes the borrower

27. Remember that in our example, C_{old} and C_{new} have equal seniority when the debt is unsecured, and C_{new} has higher seniority when it is secured. It should be noted, though, that our example does not show that it is optimal to issue new debt that has the senior-most claim against *all* of the firm's assets. Rather, the optimal new debt in the example is a prior claim against a subset of the assets and *no* claim against the remaining assets.

better off, but it even lowers the interest rate on the initial debt (C_{old} in our example). In our example, the interest rate on the loan provided by C_{old} is 57.143 percent when the issuance of debt of higher seniority in the future with respect to *any* asset is prohibited, and it is 52.86 percent when such issuance is permitted. The reason for this, of course, is that the ability to issue secured debt in the future resolves the underinvestment problem of debt.

Inadequate Effort Supply: Another moral hazard is that the borrower may expend insufficient effort in managing the firm when its assets are highly leveraged. Collateral can help to resolve this moral hazard problem, too. The following example uses outside collateral to illustrate the point.

Example 5.6 Consider an entrepreneur, Mr. David Barnes, who borrows \$100 at $t = 0$ (the start of the period) and invests the loan in a project that will pay off at $t = 1$ an amount \$300 in the successful state (state S_1) and nothing in the unsuccessful state (state S_2) for his start-up firm, Barnes Manufacturing. The probability of S_1 is p(e), where e is Mr. Barnes' effort in managing the project. Mr. Barnes can choose one of two effort levels: high (h) or low (ℓ). Mr. Barnes sustains a personal cost of \$40 to expend h and nothing if ℓ is chosen. Assume $p(h) = 0.8$ and $p(\ell) = 0.6$. Mr. Barnes has collateral available, but collateral worth \$1 to him is worth 90 cents to the bank. Assume that the bank cannot observe Mr. Barnes' choice of effort. The riskless interest rate is 10 percent. Compute the optimal loan contract.

Solution We want to show in this example that Mr. Barnes will work harder if the bank has loaned him \$100 with a secured debt contract. We will proceed in four steps. First, we will assume that the bank is restricted to offering an unsecured loan. We show that it is not a Nash equilibrium for Mr. Barnes to choose $e = h$. Second, continuing with the unsecured debt assumption, we show that it is a Nash equilibrium for Mr. Barnes to choose $e = \ell$, and for the bank to price its loan accordingly. Third, we introduce collateral and solve for the amount that makes Mr. Barnes indifferent between ℓ and h. We find that with this level of collateral it is indeed a Nash equilibrium for Mr. Barnes to choose h. Finally, in the fourth step, we check that Mr. Barnes himself is better off with secured debt, which serves as a precommitment that he will work harder.

Step 1 Suppose first that the bank restricts itself to offering an unsecured loan. If the bank assumes that Mr. Barnes will choose $e = h$, then the interest rate, r_h^u, it should charge on this unsecured loan to just break even satisfies

$$[0.8 \times (1 + r_h^u) \times 100/[1 + 0.10] = 100, \tag{5.11}$$

which yields $r_h^u = 37.5$ percent. To check if this is a Nash equilibrium, we need to ask whether Mr. Barnes, faced with this loan contract, will indeed choose $e = h$. Mr. Barnes' expected payoff with $e = h$ is

$$0.8 \times (300 - 137.5) - 40 = 90,$$

whereas his expected payoff with $e = \ell$ is $0.6 \times (300 - 137.5) = 97.5$. Thus, this is not a Nash equilibrium since Mr. Barnes prefers $e = \ell$.

Step 2 It is, however, a Nash equilibrium for the bank to assume that Mr. Barnes will choose $e = \ell$, and price the unsecured loan accordingly. The loan interest rate, r_ℓ^u must satisfy

$$[0.6 \times (1 + r_\ell^u) \times 100]/[1.10] = 100, \qquad [5.12]$$

which yields $r_\ell^u = 83.33$ percent. Mr. Barnes' expected payoff with $e = h$ is $0.8 \times (300 - 183.33) - 40 = 53.34$. His expected payoff with $e = \ell$ is $0.6 \times (300 - 183.33) = 70.00$. Thus, it is a Nash equilibrium for the bank to price its unsecured loan assuming that Mr. Barnes will choose $e = \ell$.

Step 3 Now let us see if we can do better by using collateral. Let C be the collateral that leaves Mr. Barnes indifferent between choosing ℓ and h. Then r_ℓ^u and C must be related by the following equation

$$0.8 \times (1 + r_h^s) \times 100 + 0.2 \times 0.9C = 110. \qquad [5.13]$$

The left-hand side of (5.13) recognizes that the bank is repaid in full if the project is successful (this has probability 0.8) and only collects the collateral if the project fails (with probability 0.2). The value of the collateral to the bank is 0.9C. Solving (5.13) gives

$$1 + r_h^s = 1.375 - 0.00225C. \qquad [5.14]$$

Now, the amount of collateral needed to leave Mr. Barnes indifferent between ℓ and h is given by

$$\begin{aligned} &0.8 \times [300 - 100 \times (1.375 - 0.00225C)] - 0.2C - 40 \\ &= 0.6 \times [300 - 100 \times (1.375 - 0.00225C)] - 0.4C \end{aligned} \qquad [5.15]$$

Note that in (5.15) we have substituted for r_h^s using (5.14). Solving (5.15) yields C = \$30.61. Using this value of C in (5.14) gives $r_h^s = 30.613$ percent. To have Mr. Barnes strictly prefer h, suppose we choose C = \$30.62. Mr. Barnes' payoff if he chooses $e = h$ is now the left-hand side of (5.15) with C = \$30.62 and $r_h^s = 30.613$ percent. It is \$89,386. If Mr. Barnes chooses $e = \ell$, his expected payoff is the right-hand side of (5.15) and is given by \$89,384. Hence, Mr. Barnes prefers to choose h, and it is a Nash equilibrium for the bank to offer this secured loan on the assumption that Mr. Barnes will choose $e = h$.

Step 4 Note that Mr. Barnes' expected payoff in the Nash equilibrium with *unsecured* debt is \$70, whereas in the Nash equilibrium with secured debt it is \$89,384 (if Mr. Barnes chooses $e = \ell$) or \$89,386 (if Mr. Barnes chooses $e = h$). Thus, Mr. Barnes is better off by taking a secured loan, even though the use of collateral is dissipative.

We have discussed the various roles of collateral. The type and amount of collateral used will depend on which of these problems is dominant.[28] As mentioned earlier, using collateral can be costly, however, because of repossession costs. Additional costs are created because the quality of collateral must be appraised prior to making the loan and then monitored regularly during the life of the loan.[29] The reason for the appraisal and monitoring is that variations in the quality of a particular type of collateral across different borrowers may be quite large. For example, when collateral consists of accounts receivable, it will be of much higher quality if it is pledged by a borrower that has receivables due from well-capitalized companies with triple A ratings than if it is pledged by a borrower with receivables due from weak credit risks. Another example is *contract receivables*,[30] *whose* risk increases with volatility in business cycles. The point is that all collateral is not the same, and the deployment of collateral has various costs associated with it. These costs must be traded off against the potential benefits of collateral in deciding how to use collateral in lending. We turn now to the last of the "five Cs" of credit.

(v) Conditions By this we mean the economic conditions that affect the borrower's ability to repay the loan. Debts are repaid from four sources: income, sale of assets, sale of stock, and borrowing from another source. All of these should be assessed in determining the desirability, price, and other terms of the loan. The borrower's ability to generate income depends on: the selling prices of its goods, costs of inputs, competition, quality of goods and services, advertising effectiveness, and quality of management. Analysis of the borrower's financial statements as well as its management should inform the bank about the borrower's ability to create income.

In the Appendix, we discuss recent trends in credit analysis among banks. These highlight the increasingly sophisticated usage of computer technology in credit information processing.

Sources of Credit Information

The information used in underwriting credit is inherently costly and of uneven quality. The banker's critical skill in credit lies in assembling the most germane information at the lowest possible cost without violating legal requirements or social norms. This means identifying novel sources of information and using standard sources in clever ways. Following is a brief description of some of the standard sources of bank credit information, but we should emphasize that standard uses of

28. Empirical evidence on the relationship between collateral and borrower risk appears in Hester (1979), Orgler (1970), Morsman (1986), Berger and Udell (1990), Boot, Thakor, and Udell (1990), and Jimenez, Salas, and Saurina (forthcoming). These studies find that large prime borrowers are less likely to be asked to pledge collateral, whereas *observably* higher risk borrowers usually receive secured loans. (This is *not* inconsistent with our analysis that, among a group of *indistinguishable* borrowers, collateral can sort by inducing lower-risk borrowers to pledge more collateral). The finding that large, well-known borrowers are asked to pledge less collateral is also plausible since informational problems are likely to be less severe for such borrowers.

29. See, for example, Clarke (1987).

30. A "contract receivable" is an amount that a contractor is due to receive upon successful future completion of a contract. It involves chattel paper that shows the associated monetary obligations. Loans secured by contract receivables are often created when building or manufacturing contractors, dealers, or retailers need working capital.

standard sources is unlikely to produce anything better than average results. The clever use of credit information is a cultivated art form that distinguishes the successful lender from the pack.

Standard credit sources can be classified as: internal and external. By internal sources we mean those within the bank, and by external sources we mean all other.

Internal Sources

(i) **Interview with Applicant** The loan interview normally establishes the uses to which the borrowed funds will be put for the loan request and the conformity of the application with the bank's loan policies. For example, the bank's policy guidelines usually stipulate a minimum equity input by the borrower, so that a violation of this guideline can be discussed with the borrower, leading perhaps to a smaller loan request. The loan interview is also used to judge intangibles related to the borrower's future repayment behavior. Moreover, it also provides the loan officer with an opportunity to advise the applicant about any additional financial information that might be needed for evaluating the application.

(ii) **Bank's Own Records** A bank normally maintains records of its depositors and borrowers. This source of information allows the bank to assess the borrower's past behavior.[31] For example, bank records will show the payment performance on previous loans, the balances carried in checking and savings accounts, and overdrawing patterns, if any. Even for applicants who have never been customers of the bank, the central file may contain some information if these applicants were solicited as potential customers.

External Sources

(i) **Borrower's Financial Statements** These are required of most borrowers. Audited statements are common requirements in commercial lending. Even in consumer lending, where loans are usually small, an applicant is normally asked to list what he/she owns, income and expenses, and outstanding debts.

(ii) **Credit Information Brokers** Information agencies or credit bureaus systematically collect financial information on potential borrowers and make it generally available at a price (recall Chapter 3). The most widely known is *Dun & Bradstreet* (D&B), which collects information on over 3 million businesses in the United States and Canada. D&B's *Business Information Report* provides information on the type of business, nature of ownership, composite credit rating, promptness with which the firm makes payment, sales, net worth, number of employees, general condition of the firm including information about its physical facilities, customer base, balance sheet information, the usual size of the firm's deposit balances, its payments record under loan agreements, and biographical information on principals. More detailed information can be found in D&B's *Key Account Report*. In *Dun's*

31. In Chapter 3, we pointed out that this may be an important advantage of banks in granting credit [see Fama(1980)].

Review, D&B also publishes information about financial ratios for a large number of industries.

Comparative financial information can also be found in the *Annual Statement Studies* published by *Robert Morris Associates*, a professional association of professional lenders. There are numerous other surveyors of credit information, specializing in consumer, business, and even governmental borrowers.

(iii) Other Banks Banks sometimes check with other banks that have had relationships with the loan applicant. They may also check with the firm's suppliers,[32] to learn how the firm pays its bills, and with the firm's customers to determine the quality of its products and the dependability of its service.

Analysis of Financial Statements

In evaluating the borrower's ability to service a loan, the bank will focus on the firm's internal sources for future generation of funds. These are: (i) net income, (ii) depreciation[33], (iii) reduction of accounts receivables, and (iv) reduction of inventories. To assess the potential of these cash flows, the bank examines the borrower's financial statements. However, financial statements are *noisy*. It is often necessary to work with audited statements that are months too old, along with unaudited interim statements that raise questions of authenticity. Even audited statements have their problems owing to the idiosyncracies of GAAP and the occasional lapses and professional compromises of auditors. These problems aside, financial statements value assets using nonmarket criteria such as book values, and income is distorted accordingly. Thus, financial statements should be interpreted with caution. An illustration is provided by the bursting of the stock market bubble in 2000 that was credited by some to a bond analyst raising questions about the credit worthiness of Amazon.com's debt based on accounting information not accurately reflecting cash flows for credit risk assessment purposes and concluding that Amazon's credit risk was higher than it seemed.[34]

Evaluation of the Balance Sheet

Assets

(a) Accounts Receivables: *Accounts receivables* are among the shortest maturity assets on the borrower's balance sheet and are typically seen as the major source of cash flows to service short-term loans. Standard analyses focus on the sizes, sources, and aging of accounts, as well as the extent to which the accounts receivables are actively managed and diversified. As with any other risky asset portfolio, diversification lowers risk. The bank may also wish to investigate the financial attributes of those who owe money to the borrower since these speak to the quality of the borrower's receivables. Credit bureaus are especially useful in

32. Another source of information about a potential borrower's suppliers is the *Credit Interchange Service* of the *National Association of Credit Management*.

33. Since depreciation is not a cash outflow but is subtracted in computing net income, it should be added back to arrive at cash flow.

34. Quite often, these issues are related to a divergence of accounting income from cash flows.

evaluating the quality of the borrower's receivables. Also, the current status or aging of receivables is a powerful indicator of their quality. For example, if a large fraction of receivables are 90 days or older and the convention is to pay in 30 days or less, the implications are transparent.

Not all borrowers need to be screened equally carefully. Relatively low-risk borrowers who may be close to qualifying for unsecured loans often fall under a "bulk" or "blanket-assignment" lending plan. For such borrowers, the bank may require only monthly borrowing-base certificates and aging or inventory listing, without maintaining active day-to-day control over collections. In the next risk category may be customers who keep good records and have a well-diversified accounts receivables portfolio. For such borrowers, the bank may impose additional reporting requirements, including detailed assignment, collection, and aging schedules. In the highest risk category are borrowers with weak balance sheets and inadequate working capital. Here the bank requires all standard reports plus copies of shipping documents, delivery receipts, and assigned invoices against which the bank will lend.[35] It is common for the bank to require such borrowers to remit collections directly to the bank in the form of checks "in kind." This is a way for the bank to exercise additional control. The bank might even mail invoices directly to the accounts in the borrower's accounts receivables portfolio, asking for payments to be made directly to the bank.[36]

(b) Contract Receivables: A borrower may be a contractor who has been engaged to perform some task in the future. Official recognition of this may appear in *chattel paper* that shows the monetary obligations of the party for whom this task is being performed. These monetary obligations are called contract receivables. Chattel paper often serves as collateral for a working capital loan. Contract receivables are riskier than accounts receivables since payment is conditional on the borrower's future performance. There is consequently a *double moral hazard*, one that the borrower may not successfully complete the contracted task and the other that the third party may not pay the borrower even if the task is successfully completed.[37] Thus, greater monitoring efforts are warranted for contracts receivables.

(c) Inventory: The age, liquidity, price stability, obsolescence, shrinkage, the adequacy of insurance coverage, the stage of processing, and the firm's method of inventory accounting are all issues in evaluating inventories.

As with any other form of collateral, the bank should be concerned about incentive effects as well as liquidation value. However, valuing partially processed inventories is difficult and a credit-analysis art form. Both raw materials and finished goods inventories are easier to value and have greater liquidity than partially processed goods. In many cases, raw material inventories have the broadest market and the lowest price volatility. As with other collateral, monitoring is crucial in that inventory stocks are constantly in flux, with potentially damaging consequences for the secured lender.

35. This procedure is called "ledgering" the accounts. See Clarke (1987).
36. This procedure is often referred to as handling borrowers on a "notification" basis.
37. With accounts receivables, you can see that only one of these two hazards is present.

(d) Fixed Assets: Normally, banks do not consider the sale of a fixed asset as a source of funds for loan repayment. However, surplus fixed assets can be occasional and strategic sources of cash flows. Whereas the main importance of fixed assets lies in their ability to *produce cash flows* and *not* in their resale value, business restructurings often generate surplus fixed assets whose expeditious sale can be value creating.

(e) Intangible Assets: These include trademarks, patents, copyrights, and goodwill. These assets are normally accorded little value by a bank because of their illiquidity and measurement errors. There are, of course, exceptions, but by and large bankers apply large discounts to such assets.

(f) Amounts Due: Banks often take a dim view of a firm's management if the firm's assets include amounts due from officers and employees. Amounts due create the suspicion of internal fraud and nepotism.

Liabilities and Net Worth

(a) Accounts Payable: The borrower's accounts payable should speak volumes to its bank. If the borrower does not pay its trade creditors timely, why should the bank expect to be treated differently? The bank should ascertain whether payables are in the form of notes since this may indicate that the firm has been denied trade credit. The bank should be similarly alarmed if the borrower has been asked by its suppliers for cash-on-delivery (COD) terms. In case the borrower owes money to its own shareholders or officers, the bank should demand explanation and may ask that such liabilities be subordinated to any bank loan. The bank should also review the amounts accrued for taxes and other expenses.

(b) Long-Term Liabilities: These consist of term loans, debentures, notes, mortgage loans, and other liabilities with maturities exceeding 1 year. The bank should be concerned with the *nature* and *maturity of* these obligations and the provisions that have been made for meeting the required payments. Their covenants may also be important for the bank considering a loan request. In particular, it is important to know whether the outstanding debt is secured and if so, which assets have already been pledged as collateral.

(c) Net Worth: The importance of equity capital to credit analysis is transparent, given our earlier discussions. However, *accounting* net worth is a particularly treacherous account because it is fraught with measurement errors. This item is the residual of assets and liabilities, with each asset and liability independently evaluated *with error*. Hence, the net worth compounds all of the errors embedded in the underlying accounts. If all assets and liabilities could be evaluated at market, the net worth should be the economic value of equity claims. However, with accounting distortions and other measurement errors, accounting net worth can be a hard-to-interpret residual.

(d) Contingent Liabilities: These are important because of their potential to become *actual* liabilities. If they do, they could seriously impair the debt-servicing capability of the borrower. Assessing the relevant probabilities and exposures may call for considerable information and sophistication. Moreover, such liabilities do not always appear in the body of the borrower's balance sheet. Even when footnote disclosures reveal the borrower's exposure (maximum liability), the present value of the liability depends also on the unspecified contingencies and probabilities.

The Income Statement

Income statement analysis complements balance sheet analysis. Bankers tend to emphasize the balance sheet in evaluating short-term loans, but devote greater attention to the income statement for longer-maturity loans. Recall that the balance sheet measures *stocks*, whereas the income statement measures *flows*. Hence, by looking at past and present income statements, the bank should be able to learn something about the degree of stability in the borrower's cash flows. Of course, in determining cash flow trends, the bank should be careful to note possible changes in the borrower's accounting practices can obfuscate.

The bank will often use both the balance sheet and the income statement in its *ratio analysis*. Key financial ratios convey information about the firm's liquidity, stability, profitability, and cash flow prospects.

Basically, there are four types of ratios: liquidity, activity (or turnover), profitability, and financial leverage.

(i) Two measures of liquidity are commonly used:
current ratio = current assets/current liabilities,
quick ratio (or acid test ratio) $= \dfrac{\text{current assets} - \text{inventories}}{\text{current liabilities}}$.
By "current" we mean a duration of less than 1 year.

(ii) Activity ratios include the following:
Inventory turnover ratio = sales/inventory.
Average collection period (in days) = receivables/sales per day.
Total assets turnover = sales/total assets.
Fixed asset turnover = sales/net fixed assets.

(iii) There are also numerous profitability ratios. These include:
Profit margin on sales = net profit after taxes/sales.
Return on total assets = net profit after taxes/total assets.
Return on net worth = net profit after taxes/net worth.

(iv) The leverage ratio is defined as total debt/total assets.

Perhaps the two most important leverage ratios used by lenders are: pretax interest coverage and total debt to EBITDA[38]. Pretax interest coverage is defined as net income from continuing operations before taxes divided by reported gross interest expense. EBITDA is earnings before interest, taxes, depreciation and amortization. Figure 5.7 shows the behavior of these ratios through time for investment-grade U.S. corporate borrowers. It shows that the credit risk of these borrowers has been declining since 2002.

It is worth emphasizing that these ratios are usually expressed in terms of accounting values. Since bankers evaluate these ratios against peers, it is useful to

38. See Sufi (2006), who empirically shows the important of total debt/EBITDA.

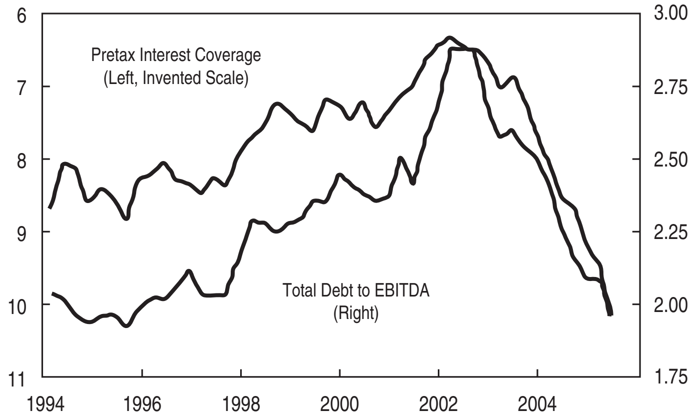


FIGURE 5.7 United States—Measures of Corporate Financial Performance for Investment Grade Corporate Borrowers (Ratio), 1994–2005
Note: Pretax interest coverage is net income from continuing operations before taxes divided by reported gross interest expense. Data are for industrial credits within the Citigroup BIG Credit Index.
Source: Citigroup.

remember that different firms may use different accounting methods. We provide a case at the end of this chapter that calls for ratio analysis as part of the credit evaluation process.

Loan Covenants

Covenants are special clauses designed to protect the bank and prohibit the borrower from taking actions that could adversely affect the likelihood of repayment. By agreeing to loan covenants that limit its actions, the borrower precommits to eschewing strategies that might expropriate wealth from the lender. The effect is to reduce the moral hazard faced by the lender and improve the terms of the loan agreement for the borrower. That is, loan covenants reduce the agency costs of debt and thereby benefit the borrower *ex ante*, and also the lender. Indeed, covenants make possible loans that would not otherwise be made at all. There is, of course, a limit to how restrictive a set of covenants the borrower will wish to accept. Restrictive covenants can make the loan reasonably safe for the lender but may deprive the borrower of valuable investment options and strategies.[39]

Loan covenants normally depend on the financial condition of the borrower, its investment opportunities, the track record of its management, and the lending

39. There may be circumstances in which restrictive loan covenants could perversely increase the likelihood of default by precluding actions the borrower could have taken to make both the bank and itself better off. For example, the purchase of new equipment by the borrower may be prohibited and yet the borrower's cash flows could be improved to such an extent by this purchase that the lender would be better off *ex post* if this covenant were relaxed. In such instances, the lender has an obvious incentive to renegotiate and relax the covenant [see Berlin and Mester (1992)]. However, if the lender is unsure of the borrower's motive for renegotiating and therefore uncertain of its potential benefit to the lender, it may refuse to renegotiate.

philosophy of the bank. Covenants are commonly classified into four kinds: affirmative covenants, restrictive clauses, negative covenants, and default provisions.

Affirmative Covenants

These are obligations imposed on the borrower. A commonly used covenant in this group is a requirement that the bank be periodically furnished with financial statements. The purpose, of course, is to permit the bank to keep track of the borrower's financial condition and enable preventive steps to be taken if trouble is indicated.

Another example is a requirement that the borrower maintain a minimum level of working capital. Banks will occasionally require the borrower to maintain a management acceptable to the bank. If management should change due to resignation, death, or other causes, the bank must approve the replacement.

Restrictive Clauses

These are designed to impose limits on the borrower's actions. A commonly used restrictive clause is one that limits the amount of dividends the borrower can pay its shareholders. The economic rationale for this covenant is transparent. A major concern for any creditor is the borrower's inclination to divert liquidity and net worth to shareholders rather than keep it within the firm to protect creditors.

It is also common for the bank to restrict salaries, bonuses, and advances to employees of the firm, as well as to limit specific types of investments such as purchases of fixed assets. The economic rationale for restrictions on investments is to protect creditors against asset substitutions that may reduce the value of the firm's debt. By purchasing a fixed asset, for example, the bank may be replacing cash on its balance sheet with an asset that will produce risky cash flows; this may increase the risk exposure of creditors.

Negative Covenants

While restrictive covenants limit certain actions, negative covenants prohibit them outright, absent the bank's consent. A common negative covenant is the *negative pledge clause*, usually found in unsecured loans. It prohibits the borrower from pledging any of its assets as security to other lenders. While the negative pledge clause is more common in unsecured loans, it is also encountered in secured loans. The banker may want to include this clause even though the bank's claim is protected with collateral because if the borrower defaults, the value of the collateral may be substantially diminished. In this case, bankruptcy law stipulates that for that portion of the bank's claim in excess of the value of the collateral, the bank has the same status as a general (unsecured) creditor. So the fewer the assets of the firm that are pledged for other loans, the greater is the share available to the bank in the event of bankruptcy.

There may also be prohibitions regarding mergers, consolidations, and sales of assets. The reason for this is that these developments can alter the firm's risk profile, possibly to the creditor's detriment. It is also common for the bank to prohibit borrowers from making loans to others or guaranteeing the debts or other performances of others. Again, the economic rationale is clear. If the borrower were to do these things, it would assume additional credit risk on its account. By prohibiting such actions, the bank protects its own claim.

These are intended to make the entire loan immediately due and payable under certain conditions. Ordinarily, even though the bank has covenants that are intended to govern the borrower's behavior, violation need not automatically empower the bank to call the loan as long as scheduled payments are being made. However, some covenants will include an *acceleration clause* that specifies *events of default*. Effectively, violation of a covenant leading to any of the events of default automatically places the loan in default and full payment becomes due immediately. This permits the bank to take more timely actions than would be possible if it had to wait until a payment was missed. Acceleration clauses are often triggered by the following:

- Failure to make timely payments.
- Inaccuracy in representations and warranties.
- Violation of covenants.
- Bankruptcy, liquidation and/or appointment of a receiver.
- Entry of a judgment in excess of a specified amount.
- Impairment of collateral, invalidity of a guarantee and/or security agreement.
- Failure to pay other indebtedness when due or to perform under related agreements
  — Cross default.
  — Cross acceleration.
- Change of management or ownership.
- Expropriation of assets.

Any of the above may be considered an event of default, in which case the loan is accelerated and will lead to either renegotiation or default. In some cases, the loan agreement provides the borrower a period of time, referred to as a cure or grace period, to correct its default. If cured, the bank is then required to continue the loan. In the case where the default is not cured, the bank may terminate the lending relationship. The bank may also set off the borrower's deposits against its obligation to repay the loan and exercise its right to foreclose on collateral and even force the borrower into receivership. The cross default provision gives the bank the right to declare an event of default when the borrower is in default on another obligation. Though banks rarely exercise the right to accelerate loan repayment, having this right substantially strengthens a lender's position.

Other Parameters of the Loan Agreement

Loan agreements have many provisions other than amount and price that must be negotiated between the bank and the borrower. Some of the more important parameters of the loan agreement are:

- A *take-down schedule*: a time table for withdrawing funds from the bank.
- An *installment schedule*: a time table for repaying the interest, other charges, and principal.

- A *compensating balance requirement*: an obligation by the borrower to maintain deposits at the lending bank. (This requirement is usually stated in terms of the *average* deposit balance but may include minima as well.)
- A *prepayment provision*: a possible penalty for repaying a loan earlier than required.

The loan agreement also may contain provisions especially tailored to a specific situation. For example:

- The borrower agrees to sell, within the next 12 months, at public auction, or by any other commercially reasonable means, a commercial property owned by the borrower located at the corner of Oak and Spring Streets in Center City.
- The borrower agrees, within 180 days, to divest himself of his interest in a partnership known as Branson Truck Lines, and to apply any and all proceeds from the sale thereof to this loan.
- The borrower agrees to obtain, as soon as possible, and to assign to the bank, $100,000 of term life insurance.

It is worth keeping in mind that covenants, no matter how elaborate, can never anticipate all contingencies and prevent all disasters. For example, a borrower could have adequate liquidity as measured by its stock of working capital, and yet its actual liquidity position may be very poor because its accounts receivables portfolio is concentrated in a few high-credit-risk accounts. No loan covenants can replace vigilant and ongoing monitoring by the bank.

Conclusion

In this chapter we have examined the bank's spot lending decision. We have seen that a loan typically is an illiquid debt contract, without an active secondary market. The distinction between bank loans and traded bonds is significant on two grounds. First, trading tends to narrow informational gaps between borrowers and lenders, so that bank loans usually have less known about them than corporate bonds. Second, banks perform valuable screening services that overcome private information problems and postlending monitoring that resolves moral hazard problems. Thus, we should expect banks to lend to borrowers about whom less is known *a priori* and to those who have a rich set of investment opportunities so that moral hazard is a concern. This suggests a way to think about which borrowers approach banks and which go to the capital market (recall Chapter 3).

We have also discussed the design of loan contracts by banks in light of the informational problems they face. We have devoted considerable attention to the role of collateral and capital in overcoming these informational problems in traditional credit analysis.

Banks use a variety of internal and external information sources in order to perform the credit analysis needed to effectively screen borrowers. We have discussed these sources to highlight the potential impact of information availability on the bank's credit decision and its loan contract design. We hope that our discussions in this chapter have convinced you that the bank's lending decision is a complex one and expertise in credit analysis, loan contract design, and postlending monitoring is a valuable resource. Hence, the uniqueness of a bank (recall Chapters 2 and 3). However, even the best experts cannot *always* effectively overcome informational problems in loan contracting. Sometimes these problems are insurmountable, and sometimes new information arrives that makes a previously negotiated loan contract inefficient. How banks deal with such situations is the subject of the next chapter.

Case Study Indiana Building Supplies, Inc.

The date is January 15, 2001. Alex Brown, vice president of the First National Bank of Bloomington (FNBB), was approached by Peter Willis, one of his loan officers who recently completed his training program at the bank after graduating with an MBA from a leading business school. Peter has been concerned about the financial ratios of one of FNBB's borrowers, Indiana Building Supplies, Inc. (IBS). The bank has installed a new software package to assist in its credit analysis, and this package monitors existing borrowers, alerting the bank to possible problems. This software package has indicated deterioration in some key financial ratios of IBS and has Peter worried about the likelihood that IBS will be able to repay the $473,000 it owes to FNBB by the due date of December 26, 2006.

Peter told Alex that he had run a special computer analysis on IBS about a month back and had noticed that some of the key financial ratios of the firm were trending downward. Peter based his assessment of IBS's ratios on the data provided in Tables 1 and 2. Not only were these ratios below the averages for the building supplies industry, but they were also at variance with the stipulations in the loan covenants negotiated between IBS and FNBB. Table 3 shows industry averages as well as loan covenant stipulations for key financial ratios for IBS. After his financial analysis, Peter contacted Bob Clemens, president of IBS, by phone and followed up with a letter providing details justifying his concerns. Clemens replied with a brief letter in which he conceded that some of the financial ratios had dipped below the levels specified in the loan covenants, but that there was no cause for alarm since the financial health of IBS was generally sound. Clemens pointed to the remarkable improvement in the firm's profit margin in 2005 relative to 2003 and 2004, and the fact that his return on net worth in 2005 was significantly above the industry average. When Peter called Clemens after receiving his reply, he explained to him that he was still concerned about the violations of ratio requirements in the covenants and wanted Clemens to send him data on the prices that IBS was charging customers for its finished goods. He also asked for (unaudited) quarterly financial statements on IBS.

Clemens seemed somewhat irritated by this request and reminded Peter that IBS had banked with FNBB for a long time and that Peter's predecessor had never been so picky with IBS even when it experienced substantially lower profit margins in 2003 and 2004. Nevertheless, he sent Peter the information he requested. When Peter analyzed this information, he found that IBS was charging higher prices than many of its competitors, especially those outside Indiana. Moreover, its quick ratio, current ratio, and its inventory turnover ratio all exhibited greater variations from quarter to quarter than the industry averages for these ratios.

IBS is a company that sells lumber products and a wide range of other building supplies in central and southern Indiana as well as in parts of Ohio and Missouri. Seasonal working capital needs as well as small capital equipment purchases have been financed primarily by loans from FNBB. IBS caters to basically two kinds of customers: local customers in southern and central Indiana and those elsewhere. Demand from the Indiana customers is somewhat erratic, but because of their strong desire to purchase from local suppliers and IBS's long-standing reputation, their demand is less sensitive to price increases than the demand of the other customers. In the past, whenever costs of raw materials have escalated, Clemens has personally visited many of his local customers and explained to them that he needed to increase his prices to keep pace with rising costs. These efforts have been successful in convincing the Indiana customers not to switch to other suppliers. Clemens has been far less successful in passing along such price increases to other customers. They usually seem to be able to locate alternative sources of supply when IBS increases its prices.

Recently, David Klinghoffer, the chief financial officer (CFO) of IBS, has been urging Clemens to confine attention to IBS's "loyal" Indiana customers, and thereby reduce the marketing costs involved in reaching out-of-state customers. In the past, Clemens was reluctant to embrace this strategy because of the erratic nature of demand from Indiana customers. When IBS was price competitive, it could always count on a predictable level of demand from its Ohio and Missouri customers. Increased competition and higher costs, however, seriously damaged IBS's profit margins in 2003 and 2004 and persuaded Clemens to raise prices in 2005 to improve profitability. Klinghoffer, who had also been advocating higher prices, pointed out to Clemens with great delight that their strategy had been a smashing success and the firm had been more profitable in 2005 than it had ever been since 2000. Thus, both Klinghoffer and Clemens were dismayed by what they viewed as "senseless pestering" by Peter Willis.

The matter has now come before Alex Brown. Peter has pointed out to Alex that FNBB has an "acceleration clause" in its loan contract that empowers it to force IBS to repay its entire loan to FNBB immediately because of the violations of covenants. Alex was hesitant to do that and decided to call Clemens. When Alex advised him of the seriousness of the situation and the possibility that the bank would insist on immediate repayment of the entire loan unless some corrective action was taken, Clemens said it was likely that IBS would need an additional 1-year loan of about $200,000 (preferably at a 10 percent interest rate) to cover the amount payable on a note that was due to another creditor in a few weeks. He also requested FNBB to advise him regarding specific steps that the bank wanted IBS to take.

After hanging up the phone with Clemens, Alex asked Peter to bring him a detailed financial analysis of IBS, along with the specific reasons why Peter was so concerned. He also asked Peter to evaluate whether IBS's request for additional credit should be approved and to recommend specific steps IBS should be asked to take if the existing loan is not accelerated and new credit is granted. Alex wants Peter to pay particular attention to the fact that the "bottom line" *does* seem to indicate that IBS has done well in 2005, which makes Peter's worry somewhat anomalous.

Questions

Imagine that you are Peter Willis. Prepare a comprehensive ratio analysis for IBS. Should the bank call back the entire loan now? Why or why not? Should FNBB be worried or is Peter just overreacting? Is it possible for IBS to generate enough cash by year-end 2006 to make full repayment to FNBB? How valid are comparisons of IBS's financial ratios to the industry average?

TABLE 1 Indiana Building Supplies, Inc. Balance Sheet
Year Ended December 31

	2000	2003	2004	2005
Cash	$100,000	120,000	90,000	70,000
Accounts receivable	400,000	480,000	600,000	600,000
Inventory	500,000	550,000	800,000	900,000
Total Current Assets	$1,000,000	1,150,000	1,490,000	1,570,000
Land and building	100,000	90,000	217,000	221,000
Machinery	150,000	260,000	202,000	179,000
Other fixed assets	85,000	66,000	27,000	15,000
Total Assets	1,335,000	1,566,000	1,936,000	1,985,000
Notes payable, bank	47,000	53,000	110,000	473,000
Accounts and notes payable	156,000	171,500	233,800	319,000
Accruals	82,000	350,500	252,200	34,300
Total Current Liabilities	285,000	575,000	596,000	826,300
Mortgage	50,000	40,000	36,000	33,000
Common stock	900,000	900,000	1,150,000*	867,000**
Retained earnings	100,000	51,000	154,000	258,700
Total Liability and Equity	1,335,000	1,566,000	1,936,000	1,985,000

* The company issued common stock in 2004.
**In 2005 the company repurchased some stock, citing the unusually low market price of its stock.

TABLE 2 Indiana Building Supplies, Inc. Income Statement

	2000	2003	2004	2005
Net sales	$5,000,000	4,400,000	$5,600,000	$4,500,000
Cost of goods sold	4,000,200	3,400,000	4,500,000	3,500,000
Gross operating profit	$ 999,800	$1,000,000	$1,100,000	$1,000,000
General administration, selling, and interest expenses	521,467	582,000	849,667	519,000
Depreciation	80,000	105,000	80,000	72,000
Miscellaneous	65,000	93,000	77,000	71,500
Net income before taxes	333,333	220,000	93,333	337,500
Taxes (40%)	133,333	88,000	37,333	135,000
Net income	$ 200,000	$ 132,000	$ 56,000	$ 202,500

TABLE 3 Indiana Building Supplies, Inc.

	Ratios Specified in Loan Covenants	**Industry Averages for 2005**
Quick ratio	$\geq$ 1.7	1.6
Current ratio	$\geq$ 2.5	2.5
Inventory turnover ratio	$\geq$ 9.00	8.5
Average collection period	NA	37 days
Fixed-asset turnover	NA	13.3
Total asset turnover	NA	3.00
Return on total assets	NA	9.5%
Return on net worth	NA	15%
Debt ratio	$\leq$ 38%	31%
Profit margin on sales	NA	3%

Notes: These figures are based on year-end figures taken from balance sheets and income statements of representative firms in the industry. These figures have been roughly constant for the past 5 years.

Review Questions

1. What are the different types of assets on a bank's balance sheet?
2. What is a "bank loan"? What are the different ways in which a bank can acquire loans?
3. Discuss the similarities and differences between loans and securities.
4. What are the major informational problems in loan contracts?
5. What is the purpose of credit analysis? Compare and contrast capital budgeting within a nonfinancial firm with credit analysis within a bank.
6. What are "the 5 Cs of credit"? What do we mean by a borrower's "character" and why is it important?
7. Can you explain intuitively why capital can resolve asset substitution moral hazard?
8. Discuss intuitively how capital can help the bank to resolve "adverse selection" problems. It would be useful to start out by explaining first what we mean by "adverse selection," and why it is a problem for the bank. Can you relate this role of capital in a bank loan contract to a venture capitalist's insistence on a minimum equity capital input by an entrepreneur seeking venture capital?
9. Please address the following questions:
   (a) What is a reverse leveraged buyout?
   (b) What are the main reasons why customers of banks become higher-quality credits after reverse LBOs?
   (c) Why are we observing such a large increase in reverse LBOs now?
10. What is the extent of secured lending among C&I loans? What are the two main types of collateral?
11. What are the costs of collateral? Why is "outside" collateral so popular despite these costs?

12. What is "underinvestment moral hazard"? Explain the intuition underlying the claim that collateral can attenuate this moral hazard. What are the implications of this for the design of bank loan covenants?
13. What is a "contract receivable"? Why is it usually more risky than an "accounts receivable"?
14. What are the main sources of credit information for banks in conducting credit analysis?
15. What is the role of ratio analysis in credit assessment? What are its limitations?
16. Overheard was the following conversation between two friends:

    Tom: I find it offensive that a bank would tell me what to do and what not to do when it makes me a loan. After all, I own the asset I'll buy with the loan because I have an *equity* stake in it. The bank is only lending me the money.

    Jack: That's nonsense, Tom! When you buy an asset with a bank loan, its the *bank* that owns the asset, and don't you forget it.

    What do *you* think? Explain your answer.
17. What are "affirmative covenants," "restrictive clauses," "negative covenants," and "default provisions"? Discuss the role of each in the design of credit contracts.
18. What are "expert systems" and what are banks attempting to achieve with them as part of credit analysis?
19. Consider a firm that has a bank loan outstanding that requires the firm to repay $900 one period hence. The firm has $300 in retained earnings that can either be paid out as a dividend to the firm's shareholders or invested in a project that will yield a single cash flow one period hence. The firm has a choice of investing in a safe project S, or a risky project R. The safe project will yield $1,000 for sure one period hence, whereas the risky project will yield $2,000 with probability 0.4 and nothing with probability 0.6. Assume that everybody is risk neutral and that the discount rate is zero. Which project has the higher *total* NPV for the firm? Which project will the firm choose, assuming that decisions are made to maximize shareholder wealth?
20. You are a bank loan officer. ABC Corporation has requested a $2.1 million loan. The corporation has $2 million in retained earnings and an existing debt obligation that calls for a repayment of $4 million one period hence. The firm has existing assets that will be worth $6 million with probability 0.7 and nothing with probability 0.3 one period hence. These are the future values of the assets in place if the firm does not make any investment at present. The firm also has the choice of investing in one of two mutually exclusive projects (A or B). Project A will yield $4 million with probability 0.7 and $2 million with probability 0.3 one period hence. Its cash flows are uncorrelated with (and in addition to) those from the assets in place. Project B will yield $13 million with probability 0.2 and nothing with probability 0.8. Its cash flows are also uncorrelated with those from the assets in place. Assume that everybody is risk neutral and that there is no discounting. Moreover, ABC's existing debt has seniority over any new bank loan. Compute ABC's project choice and your pricing of the bank loan in two cases: (i) ABC has $2 million in retained earnings that will be kept within the firm for one period, (ii) ABC has already announced that the retained earnings will be paid out as dividends right now and hence unavailable to augment ABC's cash flows one period

hence. Assume that your bank's cost of funds is zero and the bank is competitive (prices the loan to earn zero expected profit).

21. Consider a firm that needs \$350 to invest in a project that will yield a single cash flow one period hence. The firm knows the probability distribution of this cash flow, but no one else does. As a banker you only know that the firm is either low risk (L) or high risk (H). If it is L, then it will yield \$500 with probability 0.8 and nothing with probability 0.2 one period hence. If it is H, it will yield \$1,500 with probability 0.6 and nothing with probability 0.4 one period hence. The firm itself knows whether it is H or L. Assume that both the principal and interest repayments on any debt are tax deductible. The corporate tax rate applicable to this firm is 0.2. There is no equity capital on the firm's books at present, but it would raise equity if needed. The firm is locked into being either L or H, but as a banker you cannot tell which type it is. Assume everybody is risk neutral and that the discount rate (and the bank's cost of funds) is zero. Also, your bank is competitive (prices loans to earn zero expected profit). Construct a scheme consisting of two different loan contracts (one requiring the borrower to finance the project partly with equity capital and the other requiring no equity) such that the firm will truthfully reveal its private information by its choice of loan contract.
22. Consider a firm that can invest \$250 right now, at $t = 0$, in a project that will yield a single cash flow one period hence, at $t = 1$. This \$250 investment will be raised by issuing unsecured debt at $t = 0$. The project will yield \$500 with probability 0.8 and nothing with probability 0.2 at $t = 1$. Immediately after the initial investment but before the end of the period (say at $t = 1/2$), the firm can purchase another asset, call it A, for \$250 also. If purchased, A will yield a sure payoff of \$300 at $t = 1$. Those who lend the firm money at $t = 0$ cannot observe at $t = 1/2$ whether the firm had this investment opportunity. Everybody is risk neutral and the riskless rate is 12 percent. If you are the banker the firm has approached for a \$250 loan at $t = 0$, compute the price of your loan in two cases: (i) the firm can finance the acquisition of asset A with unsecured debt or not at all, and (ii) the firm can finance the acquisition of asset A with debt secured by the asset in question. Assume that in case (i), your bank (the initial lender) will have the same seniority as the new (unsecured) creditors who supply funds to purchase A. Your bank is competitive in loan pricing.
23. Given below is an excerpt from "A Friendly Conversation." Critique it.

Butterworth: I'll let that pass because I want to address your question, Mike. You know over 70 percent of business loans are secured, and collateral has some really beneficial incentive effects from the bank's standpoint. Moreover, it permits the bank to engage in creative loan-contract design that helps to resolve some thorny informational problems. It also leads to improved bank monitoring of borrowers, which is a key function associated with both secured and unsecured lending. To make a really long story short, I think that business lending is a key component of banks' activities. If regulation discourages this, then I think we'll have seriously weakened the financial intermediation process.

Moderator: If the role of banks in business lending were to diminish, what sort of losses to society do you foresee, Beth?

Butterworth: That's my favorite topic, Mike, so we could be here all night if I get going. But just briefly, I think that in the process of originating these loans, designing loan contracts, structuring covenants, including the crafting of collateral requirements, monitoring, and the restructuring of loans for borrowers in financial distress, banks have developed considerable expertise. It would be a shame if the financial system evolved in such a way that these skills would need to be relearned by others.

24. What is the "lending function" and how can it be decomposed? What is the usefulness of the decomposition?

Appendix 5.1 Trends in Credit Analysis

Banks are becoming increasingly sophisticated in credit analysis, relying more on computer-based statistical analysis of borrower attributes to determine the level of risk inherent in a particular loan. We will discuss two recent examples.

Illustration 1: Mellon Bank has installed computer software called the *Zeta Credit Scoring System* to analyze risk for private and commercial corporate clients.[1] This software has been developed by Zeta Services, Inc., Hoboken, N.J., which analyzes the financial condition of about 4,000 publicly owned firms and publishes quarterly reports for bankers. Mellon has begun using the Zeta Risk Control System both for assessing the credit risk of potential private loan customers and for monitoring existing borrowers. The system is also used by the Royal Bank of Canada.

The program produces a credit score that represents the probability that a company will stay in business and service its debt. Many banks, like Mellon, do not rely exclusively on one credit assessment. For example, Mellon has its own internally developed credit scoring system that evaluates loans. It then compares its own ratings to those yielded by the Zeta scoring system and devotes special attention to loans for which the two evaluations are strikingly different. Other banks may rely additionally on credit rating issued by the rating agencies like Moody's and Standard and Poor's. The objective, of course, is to improve the management of credit risk.

These credit scoring systems are essentially predictive models based on discriminant analysis. The purpose is to look at the data on numerous past borrowers and determine a relatively parsimonious set of variables that could have most accurately predicted which of these borrowers would default. For example, Altman (1968) provided the following formula

$$Z = 0.012X_1 + 0.014X_2 + 0.033X_3 + 0.006X_4 + 0.999X_5 \qquad [A.1]$$

where X_1 = working capital/total assets (in percentage),
X_2 = retained earnings/total assets (in percentage),
X_3 = earnings before interest and taxes/total assets (in percentage),
X_4 = market value of equity/book value of total debt (in percentage),
X_5 = sales/total assets (actual number).

1. See Gullo (1990).

Altman suggests that a Z value below 2.68 means that there is a high likelihood that the firm will go bankrupt. Since this early scoring model, numerous variants have appeared, but the idea is the same.

Illustration 2: Security Pacific Corporation has adopted a technology developed by the Department of Defense, called "neural networking." It is a branch of artificial intelligence that attempts to recreate the process by which the human brain learns.[2] The purpose of the program is to analyze risks in different types of loans. It is claimed that the neural network and the "expert systems" (a better-known branch of artificial intelligence) can solve problems that traditional number-crunching computer systems cannot.

Expert systems solve problems by utilizing the knowledge of experts in the form of "what if" statements. A neural network, on the other hand, solves problems without depending on the programmed knowledge of experts. The program is designed to "learn" and change the weights on different variables—that lead to a credit score—by detecting patterns. Neural networks are patterned on the neural connections in the brain. Their ability to learn and adapt makes neural networks appropriate for problems involving behavioral scoring and risk analysis. For example, suppose a neural network has been asked to analyze consumer mortgage loan applications. Then it will examine each variable and compare it with those in previous applications. Since the computer knows which of these previous applications were approved by the bank and which variables were weighted more heavily than others, it can compare a new application with its record of past applications and recommend a decision.

Expert systems first became popular in the mid-1980s, but as of this date only about half of the largest banks—which tend to be pioneers in the adoption of new technology—are using them. Neural networks are an even more recent adoption. Apart from Security Pacific, some other banks that are using this technology are Chase Manhattan Corporation, Manufacturers Hanover Trust Company, and Citigroup.

Limitations of Credit Scoring Models

While the use of computerized credit scoring models has grown significantly, these models are not without their shortcomings. A key shortcoming is that the estimates used in these models are based on data drawn solely from *extended* loans. Thus, these estimates suffer from *selection bias*.[3] Alternative approaches include those that rely on estimates derived from data that also include the characteristics of *rejected* applicants.[4]

References

Altman, Ed, "Financial Ratios, Discriminant Analysis and Prediction of Corporate Bankruptcy," *Journal of Finance* 23, September 1968, 589–609.

2. See Layne (1990).
3. A thoughtful review of credit scoring models is provided by Hand (2001).
4. This is known as a process of "reject inference." See Kiefer and Larson (2004).

Ausubel, Lawrence, "The Failure of Competition in the Credit Card Market," Banking Research Center Working Paper No. 153, KGSM, Northwestern University, October 1990.

Bassett, William F., and Thomas F. Brady, "The Economic Performance of Small Banks, 1985–2000," *Federal Reserve Bulletin*, November 2001, 719–728.

Berger, Allen N., and Gregory F. Udell, "Collateral, Loan Quality, and Bank Risk," *Journal of Monetary Economics* 25, 1990, 21–42.

Berlin, Mitchell, and Loretta Mester, "Debt Covenants and Renegotiation," *Journal of Financial Intermediation* 2–3, June 1992, 95–133.

Besanko, David, and Anjan V. Thakor, "Collateral and Rationing: Sorting Equilibria in Monopolistic and Competitive Credit Markets," *International Economic Review* 28, October 1987b, 671–689.

———, "Competitive Equilibria in the Credit Market Under Asymmetric Information," *Journal of Economic Theory* 42, June 1987a, 167–182.

Bester, Helmut, "Screening vs. Rationing in Credit Markets with Imperfect Information," *American Economic Review* 75, 1985, 850–855.

Boot, Arnoud, Anjan V. Thakor, and Gregory F. Udell, "Secured Lending and Default Risk: Equilibrium Analysis and Monetary Policy Implications," *The Economic Journal* 101–406, May 1991, 458–472.

Boyd, John, and Mark Gertler, "U.S. Commercial Banking: Trends, Cycles, and Policy," *NBER Macroeconomics Annual*, 1993.

Brickley, James, "Empirical Research on CEO Turnover and Firm-Performance: A Discussion," *Journal of Accounting and Economics* 36, 2003, 227–233.

Chan, Yuk-Shee, and Anjan V. Thakor, "Collateral and Competitive Equilibria with Moral Hazard and Private Information," *Journal of Finance* 42, June 1987, 345–364.

Clarke, Peter S., "Collateral Lessons," *ABA Banking Journal*, November 1987, 68–70.

Diamond, Douglas, "Reputation Acquisition in Debt Markets," *Journal of Political Economy* 97, August 1989, 828–862.

Fama, Eugene, "What's Different about Banks?" *Journal of Monetary Economics* 15, 1985, 29–39.

Green, Richard, "Investment Incentives, Debt, and Warrants," *Journal of Financial Economics* 13, March 1984, 115–136.

Gullo, Karen, "Mellon Adds Credit-Score System," *American Banker*, March 7, 1990.

Hall, John, and Timothy J. Yeager, "Does 'Relationships Banking' Protect Small Banks from Economic Downturns?" *The Regional Economist*, Federal Reserve Bank of St. Louis, April 2002.

Hand, D.J., "Modelling Consumer Credit Risk," IMA *Journal of Management Mathematics* 12, 2001, 139–155.

Hester, Donald, "Customer Relationships and Terms of Loans: Evidence from a Pilot Survey," *Journal of Money, Credit and Banking* 11, 1979, 349–357.

Huber, Stephen K., *Bank Officer's Handbook of Government Regulation*, Second Edition, Warren, Gorham & Lamont, 1989.

Jackson, T.H., and A.T. Kronman, "Secured Financing and Priorities Among Creditors," *Yale Law Review* 88, 1979, 11–43.

Jimenez, Gabriel, Vicente Salas and Jesus Saurina, "The Determinants of Collateral," forthcoming, *Journal of Financial Economics.*

Kiefer, Nicholes, M., and C. Erick Larson, "Specification and Information Issues in Credit Scoring," *Economic and Policy Analysis Working Paper 2004*, Comptroller of the Currency, December 2004.

Kulkosky, Edward, "21% of Banking Assets Now Tied to Mortgages," *American Banker*, January 21, 1994.

Layne, Richard, "Security to Try Neural Networking," *American Banker*, July 3, 1990.

Leland, Hayne, and David Pyle, "Information Asymmetries, Financial Structure, and Financial Intermediaries," *Journal of Finance* 32, May 1977, 371–387.

Morsman, E. Jr., "Commercial Loan Structuring," *Journal of Commercial Bank Lending* 68–10, 1986, 2–20.

Myers, Stewart, "Determinants of Corporate Borrowing," *Journal of Financial Economics 5*, 1977, 147–175.

Orgler, Yair, "A Credit Scoring Model for Commercial Loans," *Journal of Money, Credit and Banking* 2, 1970, 435–445.

Rose, Sanford, "Why Banks Make So Many Bad Loans," *American Banker*, June 19, 1990.

Ross, Stephen, "The Determination of Financial Structure: The Incentive-Signalling Approach," *Bell Journal of Economics and Management Science*, Spring 1977, 23–40.

Shah, Salman, and Anjan V. Thakor, "Optimal Capital Structure and Project Financing," *Journal of Economic Theory* 42, August 1987, 209–243.

Sheshunoff, Alex, "Best of Times or Worst of Times?" *ABA Banking Journal*, July 1988, 25–37.

Smith, Clifford W., and Jerold B. Warner, "On Financial Contracting: An Analysis of Bond Covenants," *Journal of Financial Economics* 7, 1979, 117–161.

Stulz, Rene M., and Herb Johnson, "An Analysis of Secured Debt," *Journal of Financial Economics* 14–4, December 1985, 501–522.

Sufi, Amir, "The Real Effects of Debt Certification: Evidence from the Introduction of Bank Loan Ratings," Working Paper University of Chicago, April 2006.

This page intentionally left blank

CHAPTER ♦ 6

Further Issues in Bank Lending

"A banker is a fellow who lends you his umbrella when the sun is shining and wants it back the minute it begins to rain."

Mark Twain

Glossary of Terms

Discount Window: A facility, often referred to as lender of last resort, where banks can borrow short term from the Federal Reserve to meet their liquidity needs, normally using Treasury securities as collateral. The interest rate charged for these advances, a tool of monetary policy, is called the "discount rate."

Open Market Operations: Purchases and sales of government securities by the Federal Reserve to adjust the legal reserves available to banks to support their deposit liabilities. Sales of government securities to banks reduce the reserves available to banks, and purchases of government securities from banks increase these reserves. This is a tool of monetary policy.

Interest Elasticity of Investment: measure of the sensitivity of demand for investment funds by corporations to changes in interest rates (their borrowing rates).

Monetary Policy: The Central Bank's (Federal Reserve's) policy with regard to the money supply and interest rates.

Reserve Requirement: The fraction of bank's deposits that must be kept as liquid assets, either vault cash, or deposits with the Federal Reserve.

CD: A certificate of deposit. This is a time deposit with a stated maturity and interest rate. It may be negotiable (marketable) or nonnegotiable (nonmarketable).

Consol Bond: A bond with an infinite maturity, that is, one that promises a perpetual coupon stream and has no principal repayment.

Credit Crunch: Precipitous reduction in the availability of credit.

Introduction

In Chapter 5 we examined informational problems in lending and how these problems are addressed through the design of loan contracts. In this chapter, we continue our discussion of loan transactions and extend it to cover a variety of issues such as the initial pricing of loans and adjustments in contractual terms that take place after the loan is made. While Chapter 5 was concerned mainly with static issues in lending, this chapter is concerned mainly with dynamic issues. We begin in the next section with a discussion of how profit margins are assessed and how loans are priced. In the section that follows, we examine the reason for possible price rigidities in loan contracts and credit rationing. The bank's optimal lending process is described in the next section. We then explore the economic incentives for banks and borrowers to develop long-term relationships. This is followed with a discussion of loan default and restructuring. A case study is presented to help illustrate the concepts.

Loan Pricing and Profit Margins: General Remarks

In this section we discuss how banks assess the profitability of loans and how these are priced. We begin our discussion with an analysis of the assessment of profit margins. This is followed by a discussion of benchmark lending rates, after which we discuss compensating balances. We conclude the section with an analysis of the link between default risk and bank profit margins.

Assessing Profit Margins

To assess the profit margin of a loan, a bank should first determine its sources of income from lending.[1] These are (a) the interest on the loan, (b) noninterest fee income on the loan, and (c) income from fees charged for services the borrower purchases due to the lending relationship. As for (b), there are many sources of noninterest fee income. These include closing fees (charged for concluding the loan agreement) and loan servicing fees. As for (c), borrowers may purchase a variety of services from banks due to the lending relationship. These include cash management services and trust services, for example. If the purchase of these services can be linked to the taking of the loan, then the net profit from the sales of these services by the bank should be attributed to the loan.

After assessing the income from the loan, the bank should compute the expenses incurred to generate that income. These expenses include processing costs, salaries,

1. See Warberg (1971) for a discussion.

TABLE 6.1 The Profit Equation

INCOME	–	EXPENSES	–	COST OF FUNDS	=	PROFIT
Loan Interest		Loan Processing Costs		Cost of Demand Deposits		
Noninterest Fee Income		Salaries		Cost of Time Deposits		
Income from Bank Services		Postage		Cost of Nondeposit Funds		
		Advertising and Marketing		Servicing Costs		
		Occupancy Costs				

postage, advertising and other marketing expenses, occupancy expenses, and other loan servicing costs. Finally, the bank should compute the costs of funding the loan. These costs include the cost of demand and time deposit and nondeposit funds supporting the loan, as well as the costs of servicing deposits. Having assessed income expenses and costs, the bank can calculate its profit on the loan as shown in Table 6.1.

Benchmark or Reference Lending Rates

Our previous discussion of profit margins did not explain how a particular loan interest rate itself should be determined. In practice, banks set the interest rates on loans by relating them to a benchmark or reference interest rate. A commonly used reference rate is the *prime interest rate*[2]. Traditionally, the prime rate was the interest rate posted by the bank for short and intermediate maturity loans for its most creditworthy customers, usually corporations with "blue-chip" credit ratings.

Nowadays, the bank's most creditworthy customers pay less than the prime. The prime is an administered rate loosely linked to market interest rates, and it tends to be more sluggish than market rates.

Determining the prime rate is one of the many decisions a bank makes in the process of managing its balance sheet. Whereas each bank sets its own prime lending rate, the behavior of competing financial institutions is a major influence. In addition, three major categories of market interest rates provide the principal inputs in the prime-rates setting process: (a) the rates on nonloan bank assets, (b) rates on bank-acquired liabilities, and (c) rates on corporate debt claims that are close substitutes for bank loans. Also, the term structure of interest rates, bankers' expectations of future interest rates, the expected growth in deposits, and the expected growth in loan rates are important in setting the prime.

Many of the bank's loan rates are indexed to the prime rate, either additively as in "prime plus" (that is, prime plus 1 percent) or multiplicatively as in "prime times" (that is, prime 1.05). Thus, a decision to alter the prime rate involves adjustments in a bank's entire schedule of business loan rates. This means that a bank must consider expected demand for all types of loans in determining its prime rate.

Later in this chapter we will discuss bank-customer relationship, a particularly important topic in view of the growing emphasis on *relationship banking*. For now, it

2. See Merris (1975) for example, another reference lending rate is the London Interbank Offer Rate (LIBOR), which is the virtually risk-free rate on short-term borrowing between banks in the London credit market. The Fed Funds rate is the analog in the U.S. market.

suffices to note that "customer relationships" are arrangements whereby a bank provides a variety of services to long-established customers, and these relationships must also be considered in setting the prime rate. Customers are typically risk averse and hence dislike frequent and unpredictable adjustments in their borrowing rates. Thus, in order to foster customer relationships, the bank may wish to smooth the prime rate in relation to market interest rate movements. The usual customer relationship includes two features that are particularly relevant to prime rate determination—compensating balance requirements and loan commitments. We will deal with loan commitments in the next chapter. Compensating balances are dealt with next.

Compensating Balances

Increased competition in banking in recent years has reduced the use of "compensating balances." Nevertheless, some banks still require minimum average deposit balances (known as compensating balances) as partial compensation for bank loans and other bank services. The bank's compensation results from not paying interest (or paying below-market interest) on compensating balances.

Compensating balances frequently are used with loan commitments or lines of credit. They can be viewed as raising the effective loan rate. Although compensating balances requirements are usually stated as percentages of the dollar amounts of credit lines, many arrangements require the deposit of additional balances when credit lines are activated or used. Nominal loan rates are quoted in terms of the loan principal. If a borrower must use a part of the loan to meet compensating balances requirements, the effective loan rate on the funds available for the borrower's use will exceed the stated rate because the borrower is paying loan interest on funds committed to remain in his deposit account. This means that a bank can increase effective loan rates by simply increasing compensating balance requirements and leaving its prime rate unchanged. In other words, given the fact that the prime rate affects the bank's entire schedule of lending rates, the bank may respond to changes in market interest rates by leaving the prime unchanged but changing nonprice loan terms—maturities, collateral requirements, or compensating balance requirements—so that effective lending rates can be *selectively altered*.[3]

The Relationship Between Lending Profit and Default Risk

How should a bank set the interest rate on loan? In the previous chapter, we made the simplifying assumption that each loan is priced to yield zero expected profit to the bank. As mentioned earlier, this is a representation of perfect competition among lenders. Such prices should only be viewed as minimal, however, since loan markets are imperfectly competitive. Thus, loans will be priced so that banks earn profits. The question is: How should the price of the loan be related to its riskiness? We will show that, because of agency problems, banks may price loans so that riskier borrowers are charged less than safer borrowers on a *risk-adjusted* basis.

3. See Sprinkle (1987).

Example 6.1 To examine this issue, imagine that banks can charge any borrower 150 basis points above the interest rate at which the bank would break even (in an expected value sense) on that borrower. This is a simple way to recognize the inertia induced by transactions costs or switching costs. That is, the bank can charge a borrower 1.5 percent above its breakeven rate before the customer will consider switching to another bank. By assumption, the bank's own borrowing cost is the riskless interest rate. Now suppose the bank has two types of borrowers who are observationally separable. One is a low-risk borrower, Safeway, Inc., and the other is a high-risk borrower, Gamble Brothers. Although the bank can distinguish between these two types, it cannot directly control what the borrower does with the bank loan. Each borrower has the choice of investing in one of two mutually exclusive, single-period projects: S and R, each of which requires a $100 investment. The cash flow probability distributions of these projects are given below ("w.p." means "with probability").

TABLE 6.2 Probability Distribution of Project Cash

Borrower Type	Cash Flow Distribution for S	Cash Flow Distribution for R
Low risk (Safeway, Inc.)	$150 for sure	$153 w.p. 0.9 and zero w.p. 0.1
High risk (Gamble Brothers)	$150 w.p. 0.8 and zero w.p. 0.2	$161 w.p. 0.5 and zero w.p. 0.5
The riskless interest rate is 5%.		

Compute the bank's expected profit on each borrower.

Solution We solve this problem in three steps. First, we examine Safeway, Inc. and ask what project the bank would like Safeway to choose. It turns out the answer is S. We then solve for the interest rate the bank can charge that will induce Safeway to choose S. Second, we examine Gamble Brothers. If the bank assumes that this borrower will choose R, then the breakeven interest rate is so high that the borrower declines the loan. We solve for the interest rate that induces Gamble Brothers to choose S. Finally, in step 3 we compute the bank's expected profit on each borrower, and find that this profit is higher on Safeway. Note that one key assumption here is that the bank is unable to directly control the borrower's project choice, so that it must attempt to influence it through its loan pricing. Another key assumption is that the markup over the breakeven interest rate that the bank can charge is constant across borrowers.

Step 1 Consider first Safeway, Inc. If the bank assumes that this borrower will choose S, then its breakeven loan interest rate is 5 percent. Since it can charge another 1.5 percent without losing this borrower, it can post a loan interest rate of 6.5 percent. We can see that if the bank charges this interest rate, Safeway's net expected payoff is

(i) $150 - 106.5 = \$43.5$ if project S is chosen and
(ii) $0.9(153 - 106.5) = \$41.85$ if project R is chosen.

Thus, the bank's assumption about Safeway's project choice is validated. Note that since the markup over the breakeven interest rate is fixed, the bank's expected profit is higher

(Continued)

the lower is the riskiness of the project that the borrower chooses. Thus, it is in the bank's interest to ensure through its loan pricing policy that the borrower chooses S rather than R. In the case of Safeway then, the bank can charge an interest rate of 6.5 percent.

Step 2 Consider now Gamble Brothers. If the bank assumes that this borrower will select R, then it must set the repayment obligation on the loan at \$210 to break even (that is, note that $[(\$210 \times 0.5)/1.05] = \$100$). But Gamble Brothers would not take a loan at those terms. If the bank assumes that Gamble Brothers will choose S, then its breakdown interest rate is 31.25 percent (that is, $[(\$131.25 \times 0.8)/1.05] = \$100$). We can verify that as long as the interest rate is no more than 31.667 percent, Gamble Brothers will prefer S to R. Thus, let us say that the bank will charge 31.66 percent.

Step 3 We can now compute the bank's net expected profit on each borrower. On Safeway, Inc., the bank earns a net profit of \$1.5 or 1.5 percent. On Gamble Brothers, the bank's net expected profit is $(\$131.66 - \$131.25) \times 0.8 = \$0.328$. That is, the bank earns a higher expected profit on the low-risk borrower than on the high-risk borrower,[1] even though it charges the latter a higher loan interest rate.

1. Empirical support for this observation is now provided by the Loan Pricing Corporation of New York. See Rose (1990).

The intuition is as follows. A high-risk borrower has riskier projects than a low-risk borrower and therefore the bank's breakeven interest rate on such borrowers is higher, that is, high-risk borrowers must be charged a relatively high interest rate even *before* the bank's profit margin is considered. Further, because their probability of repaying the loan is lower, such borrowers must be charged a higher *nominal* interest rate premium over the breakeven rate for the bank to earn a *given* profit. However, as our example shows, the higher the interest rate charged by the bank, the greater is the borrower's desire to switch to a riskier project. This is a general result. It is intuitive because a high repayment obligation means that even if the project succeeds, the borrower's net payoff after repaying the bank is relatively low, and perhaps even negative. This makes it more attractive for the borrower to gamble on projects that yield larger payoffs if they are successful but have lower success probabilities. The bank rationally anticipates such behavior by the borrower. It realizes that to earn the same expected profit on the high-risk borrower that it does on the low-risk borrower, it will have to charge the high-risk borrower such a high interest rate that the borrower would be induced to choose greater risk than the bank would like. In other words, the bank has less room to earn profits on the high-risk borrower because increases in interest rates discourage such borrowers from choosing the desired relatively safe investments.

The management implication is obvious. Banks may wish to refocus their attention on the low-risk, low-spread borrowers. Deposit insurance has distorted these incentives and induced banks to pursue riskier investments than would otherwise be optimal. Moreover, to the extent that riskier borrowers are less well known, the intermediation rents that banks can earn from servicing these borrowers may also be greater. This too creates incentives for banks to pursue riskier borrowers. It turns out that the incentive effects of interest rates influence the *overall* allocation of credit, not just the pricing of loans. This is an issue we examine in the section on credit rationing.

The Mathematics of Loan Pricing

Having provided the basic background for loan pricing, we now develop the mathematics behind how loan processes are determined. It turns out that bank loan pricing has a close relationship to the principles of capital budgeting used by nonfinancial firms.

The Basic Components in the Loan Pricing Equation

The bank would like to set the price of the loan so as to have NPV ≥ 0 to the bank. To ensure NPV ≥ 0, the expected loan revenues must exceed the bank's "cost of funds" plus the "institutional costs" of making the loan, that is.

Expected loan interest revenue:

$$\begin{aligned}&\geq \text{Institutional cost of loan}\\&+[\text{amount of debt financing in the loan} * \text{cost of debt}]\\&+[\text{amount of equity financing in the loan} * \text{cost of equity}].\end{aligned}$$

Since expected loan revenue:

$$= [\text{loan interest rate} * \text{size of loan}] - \text{expected loss on the loan},$$

we can write:

$$\begin{aligned}\text{Loan Interest Rate} \geq &\left[\frac{\text{Institutional Costs}}{\text{Loan Size}}\right]\\&+\left[\frac{\text{Expected Loss on Loan}}{\text{Loan Size}}\right]\\&+\left[\frac{\text{Debt Financing in Loan}}{\text{Loan Size}} * \text{Bank's Cost of Debt}\right]\\&+\left[\frac{\text{Equity Financing in Loan}}{\text{Loan Size}} * \text{Bank's Cost of Equity}\right].\end{aligned} \qquad [6.1]$$

Institutional Costs

The institutional costs of making a loan are the direct cost of monitoring the loan and the collateral, the direct costs of screening the applicant, and the allocated overhead costs. Included in the allocated overhead costs are the costs of using property, plant and equipment, and the costs of regulation and management.

There are various empirical estimates of institutional costs that are available for United States banks. For example, Oliver, Wyman & Company estimates them to be about 150 basis points, whereas McKinsey & Company estimates them to be approximately 250 basis points. Of course, this cost will vary depending on the size of the bank, the market in which it operates, the existing regulations and the type of loan.

Expected Loss on a Loan

The formula for this is:

Bank's expected loss on a loan = probability of default × the expected loss *given* default.

Figure 6.1 shows how each component of the expected loss on a loan behaves as a function of the value of the borrower's asset given that the borrowing is secured with the project financed by the loan.

In practice, banks often use a "recovery rate" of 30 percent, implying an expected loss given default of 70 percent. Oliver, Wyman & Company estimates an average probability of default for mid-market lending of about 1.2 percent.

Many banks now use the borrower's credit rating to estimate probabilities (this is also consistent with the approach in the Basel II Capital Requirements that we will discuss in a later chapter). Moody's KMV, a division of Moody's Corporation estimates ranges of default probabilities based on credit ratings as follows:

AA/Aaa = 0.02% – 0.03%
AA/Aa = 0.03% – 0.10%
A = 0.10% – 0.24%
BBB/Baa = 0.24% – 0.58%
BB/Ba = 0.58% – 1.19%

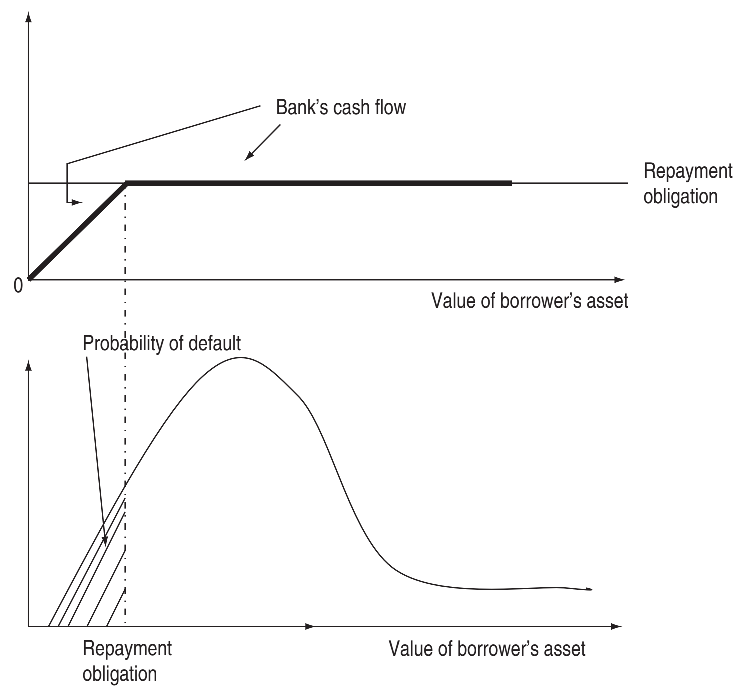


FIGURE 6.1 The Bank's Expected Loss on a Loan

The Capital Structure Supporting a Loan

Just like a nonfinancial firm finances its assets with a mixture of debt and equity, so does a bank finance its loan with a mixture of debt and equity. How does a bank determine the mix?

Here we use a well-known result from corporate finance, namely that firms with more volatile cash flows and higher assets betas (greater systematic risk) use more equity in their capital structures. Similarly, a bank will use more equity capital in its financing of a loan that has higher potential for cash flow volatility and thus higher risk. In practice, a bank will create numerous loan categories and decide which category a particular loan belongs to. Each loan category will have a hypothetical capital structure, and categories associated with more risk will have more capital allocated to them.

The Required Rate of Return on the Bank's Debt and Equity Capital

The pretax cost of the bank's debt is simply the average cost of all of the bank's debt. This includes the costs of various types of insured and uninsured deposits, the cost of various forms of nondeposit short-term borrowings like advances, and the cost of subordinated debt. Then:

$$\text{Cost of debt} = \text{average pretax cost of debt} * [1 - \text{T}] \qquad [6.2]$$

where T is the bank's effective tax rate.

What determines the cost of the bank's equity capital? This is the minimum expected rate of return that the bank's shareholders demand, given the risk in their investment. Now bank assets are unique because they are primarily *debt claims*. This means that the bank's payoff on a loan is fixed *unless* default occurs. In computing the risk of default, the bank must assess the default risk of a single asset as well as the default risk that a single asset adds to a diversified portfolio.

Default Risk of a Single Loan: Suppose a bank is considering lending to a firm. If it makes the loan, the firm will have approximately \$75 million of debt due in one year and an expected market value of assets of \$150 million in one year. The standard deviation of the firm's assets is assumed to be 17 percent. See Figure 6.2

What *Figure 6.2* gives is a single number representing the probability of default. It is what is *expected*. It does not tell us the bank's actual losses, which are random variables with probabilities associated with them. Thus, we need to characterize the *distribution* of losses as well. For each loan in the portfolio, we can characterize the probability of losses using: (i) the mean loss (expected loss) and the (ii) loss volatility.

To see this with an example, suppose a bank has made a loan to a firm on an island where it rains on one side or the other in a given year, but never on both sides. The probability of rain on any given side is 0.5. Assume that the loan repayment is \$1 million and the loss given default is 100 percent. In this case, the bank's expected loss $= 0.5 * \$1$ million $= \$0.5$ million. The loan loss volatility = standard deviation of loan loss:

$$\sqrt{0.5[\$1 \text{ million} - \$0.5 \text{ million}]^2 + 0.5\,[0 - \$0.5 \text{ million}]^2} = \$0.5 \text{ million.}$$

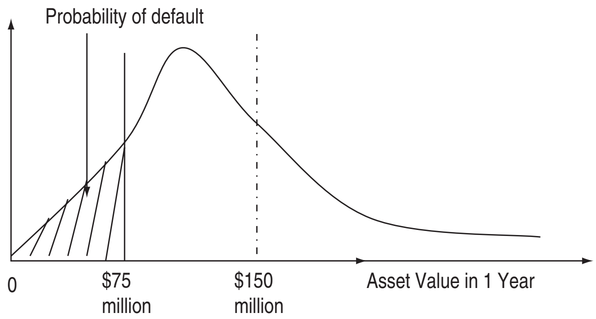


FIGURE 6.2 Distribution of Firm's Asset Value

Default Risk of a Loan Portfolio: Now let us consider the effect of forming loan portfolios. Just as with expected returns, the expected loss of a loan portfolio is the weighted average of the individual expected loan losses, adjusted to take into account portfolio diversification effects. To see how diversification affects the loan loss volatility of the portfolio, suppose that the bank now makes two loans, one to a farm on one side of the island and another to a farm on the other side. Assume each loan is \$0.5 million, so the total amount loaned out is \$1 million. What is now the distribution of losses in the loan portfolio?

Note first that the bank's expected loan loss is still \$0.5 million (the sum of the expected loan losses on the two loans, each of which is 0.5*\$0.5 million = \$0.25 million). The loss volatility on each loan is

$$\sqrt{0.5[\$0.5 \text{ million} - \$0.25 \text{ million}]^2 + 0.5\ [0 - \$0.25 \text{ million}]^2} = \$0.25 \text{ million}.$$

Recognizing that each loan has a weight of 0.5 in the portfolio and that the two loans are perfectly negatively correlated, we can use [1.7] to obtain the portfolio loan loss volatility as:

$$\sqrt{(0.5)^2(\$0.25 \text{ million})^2 + (0.5)^2(\$0.25 \text{ million})^2 - 2(0.5)(0.5)(\$0.25 \text{ million})\$0.25 \text{ million}} = 0.$$

Thus, portfolio diversification eliminates loan loss volatility in this case.

This means that the amount of equity capital supporting a loan depends on the characteristics of the portfolio that the loan belongs to. When the bank adds a loan to an existing portfolio, it computes the impact of this additional loan on the loan loss volatility of the portfolio in order to compute the incremental loss volatility due to the loan and consequently the equity capital needed to support the loan.

Distribution of Portfolio Losses

The distribution of portfolio losses is *not* normal. In practice, the distribution is very skewed. As Figure 6.3 below shows, there is a high probability of "small" (less than expected) losses, and a small (but positive) probability of extremely large losses.

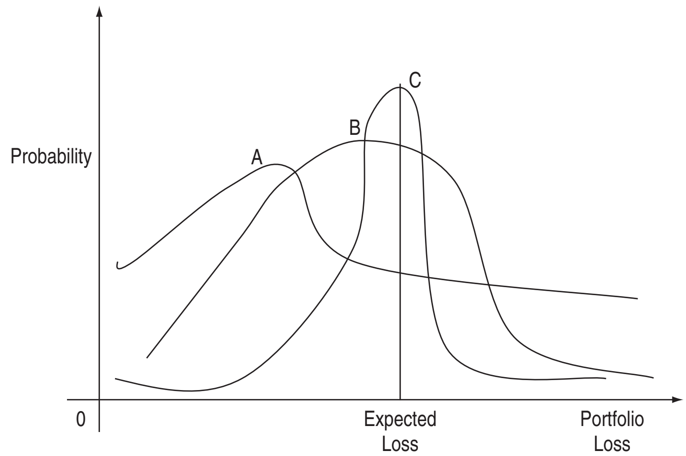


FIGURE 6.3 Distribution of Portfolio Losses and the Effect of Diversification

In Figure 6.3, curve A represents the distribution of portfolio losses when the portfolio is not very well diversified. The variance of losses and, hence, the loan loss volatility is quite high. Moreover, the distribution is skewed in that mean lies to the right of the peak of the distribution, i.e., there is a relatively high probability of losses that are smaller than the expected loss. As the portfolio becomes better diversified, we move to Curve B, which as a distribution diversification makes the distribution with a lower loan loss volatility. Further diversification makes the distribution look like Curve C, which is beginning to concentrate most of the high-probability outcomes around the mean or the expected loss. In the limit, as the portfiolio becomes perfectly diversified, as in the case of the portfolio of loans to the two firms considered earlier, the distribution collapses to a single point represented by the expected loss, i.e., all loan loss volatility is eliminated.

Recap and Summary

Once the bank has estimated the equity capital to be committed to a loan, it can use (6.1) to determine the minimum loan interest rate.[4] The actual interest rate will depend on market conditions; the greater the bank's monopoly power in a given market, the greater will be the (positive) spread between the loan interest rate and the minimum rate given by (6.1). A summary of the loan interest rate determination is given in Figure 6.4.

Some of the important additional considerations are that loan commitments should be included in the analysis. Moreover, it should be recognized that covenants in the loan contract reduce the risk of a new loan, highly ''concentrated'' (say in a particular industry as loans to firms of similar size) portfolios should require more capital.

4. The equity cost of capital used in (6.1) can either be just the bank's overall equity cost of capital or it can be a loan-specific cost of capital, adjusted to account for the riskiness of the loan relative to the whole bank.

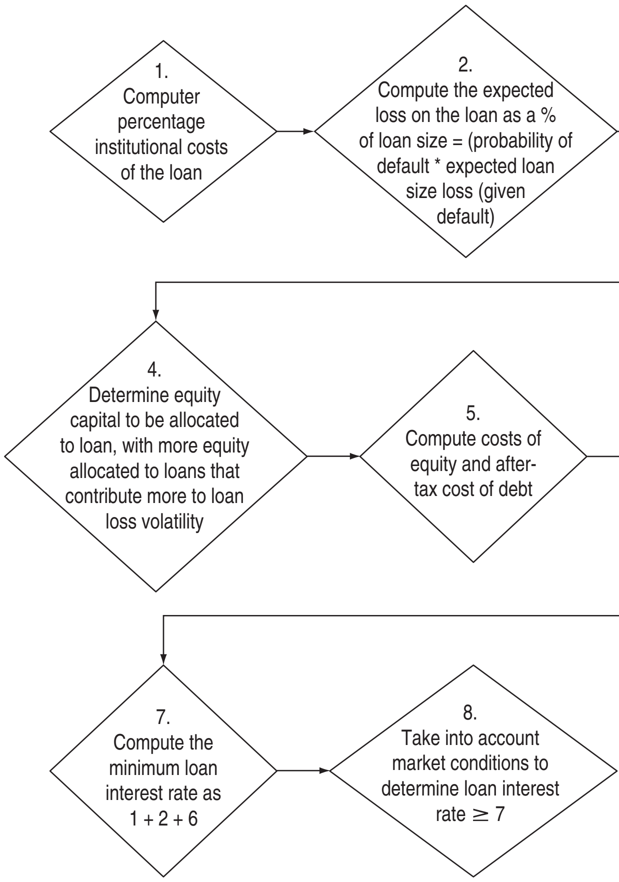


FIGURE 6.4 Loan Interest Rate Determination

Credit Rationing

Credit rationing is defined as a situation in which a lender refuses to extend credit to a borrower at the price *posted by the lender* for that borrower class. Credit rationing is *not* a phenomenon whereby a potential borrower refuses to accept credit because the price is "unfair" or too high. The essential point is that credit is denied at a price selected by the lender itself. Even if the borrower offers a higher interest rate than that asked for by the lender, a loan is refused by the lender.

Credit rationing is a puzzling practice.[5] When credit is rationed, there is an unsatisfied demand for credit at the price posted by the bank, that is, credit demand exceeds supply at that price. Conventional economic theory, or just plain common sense, suggests that the bank could increase its profits by increasing the price of credit. If the supply function for credit is upward sloping and the demand function is downward sloping, as shown in Figure 6.5, then this should bring about the usual equilibrium in which demand and supply are equated. Since the bank is supplying

5. Included in credit rationing is the practice of "redlining," which involves the lender refusing to extend the credit based on considerations of race, gender, and so on. This is illegal and is not the focus of our discussion.

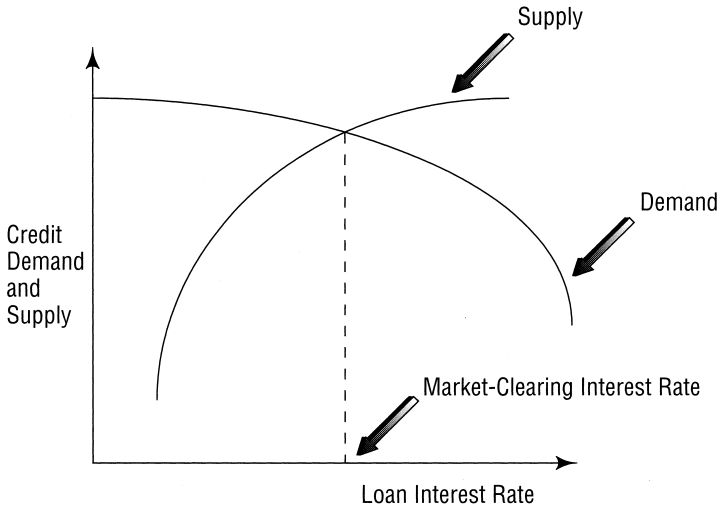


FIGURE 6.5 The Demand and Supply for Credit

more credit and at a higher price, its profit should be greater. Thus it seems irrational for profit-maximizing banks to ration credit.[6] Is it?

While it is conceivable that banks forgo profitable lending opportunities, it seems implausible. We thus ask whether it is rational for a profit-maximizing bank to ration credit.

Why Should We Be Interested in Credit Rationing?

It is believed that a fall in the money supply restricts spending. This could happen even if the fall in the money supply caused only a small increase in interest rates, or if spending is not curtailed by an interest rate increase. The reason is that a fall in the money supply would leave banks with less to lend, forcing them to reduce their lending, even if customers did not reduce their loan demand. Thus, spending was viewed as being constrained by the availability of credit to banks, and this credit was allocated to customers through nonprice means such as credit rationing. This argument, popularly known as the "availability doctrine," suggested an alternative transmission channel for monetary policy that was based in an important way on the monetary policy argument.

There are two reasons why we should be interested in studying credit rationing in connection with monetary policy. First, with credit rationing, monetary policy can be effective influencing aggregate investment by corporations even with little variation in interest rates. That is, if the Federal Reserve feels that inflationary pressures need to be abated by curtailing spending, it could cause a slowdown of the economy without major changes in interest rates. This could be achieved by reducing the liquidity of banks, which in turn could lead to reduced bank lending due to credit rationing, even if investment demand by corporations was unchanged. Thus, the effectiveness of monetary policy would have *not* been empirically documented. An important implication of this is that in the presence of credit rationing, the monetary policy options of inducing increased interest rates through a higher discount window borrowing rate

6. Samuelson (1952) was the first to suggest this.

and of reducing the amount of credit available through open market operations (bond sales) are not necessarily equivalent. Credit can be reduced even if investment demand is insensitive to monetary policy manipulations.

Second, it has been empirically found that a more stringent monetary policy does not affect all borrowers equally. Thus, if credit rationing is better understood with respect to the identities of those who are rationed, we may be able to better predict the effects of a restrictive monetary policy.[7]

Why Is There Credit Rationing?

In order to understand why a profit-maximizing bank might ration credit, we need to examine the conditions under which it would not be optimal for the bank to increase its loan interest rate when faced with excess demand for credit. It is difficult to see why banks would do this if they had as much information as the borrower. If the bank was perfectly informed, it could always set an appropriate risk-adjusted price and lend accordingly.

However, in a world of asymmetric information, credit rationing can be an optimal strategy for a profit-maximizing bank. The explanation turns upon two types of information hurdles.[8] First, a bank may not be able to distinguish perfectly between borrowers with different credit risks, even after it has analyzed each borrower's financial information. This is called the *precontract private information* problem. Even if the bank knows the *average* riskiness of borrowers within a given risk classification, it may not be able to identify individual risks [recall the Akerlof (1970) discussion in Chapter 1]. The bank will, therefore, charge a common price to all within the risk class, so that some borrowers are subsidizing others. A second problem is that the bank may not be able to completely control the borrower's actions. The borrower may thus be able to increase project risk, either through its choice of projects or through its expenditure of effort, without detection by the bank.

Now imagine that a loan interest rate is announced by the bank for a particular risk class, and at that interest rate there is an excess demand for loans by borrowers in that risk class. What would happen if the bank chose to increase the loan interest rate? One possibility is *adverse selection*. Safer borrowers within the given risk classification may be unwilling to borrow at the higher interest rate, so that the mix of borrowers within the pool becomes riskier. If this happens, the bank's expected profit could actually be *lower* at the higher interest rate; we provide a simple numerical example below to illustrate. A second possibility is that an increase in the loan interest rate could *worsen* the moral hazard problem. That is, those borrowers within the pool who have some latitude in their investment decisions may choose riskier projects at the higher loan interest rate. This again could mean a lower expected profit for the bank at the higher loan interest rate. Thus, the bank may conclude that increasing the loan interest rate is not worthwhile since its expected profit is maximized at an interest rate at which credit demand exceeds supply.[9] Figure 6.6 depicts this graphically.

7. Some evidence suggesting rationing is provided in Jaffee and Modigliani (1969).

8. What follows is an adaptation of Stiglitz and Weiss (1981).

9. That is, suppose r is the loan interest rate, C is the bank's per dollar cost of funds and θ is the repayment probability. Then the bank's expected return per dollar loaned is $\rho = [1 + r]\theta - C$. The point is that θ cannot be taken as being unaffected by r. As r is raised, θ falls. Assuming that θ is a decreasing and concave function of r (that is, $\partial\theta/\partial r < 0$, $\partial^2\theta/\partial r^2 < 0$), we see that the function $\rho(r) = [1 + r]\theta(r) - C$ attains a unique maximum with respect to r.

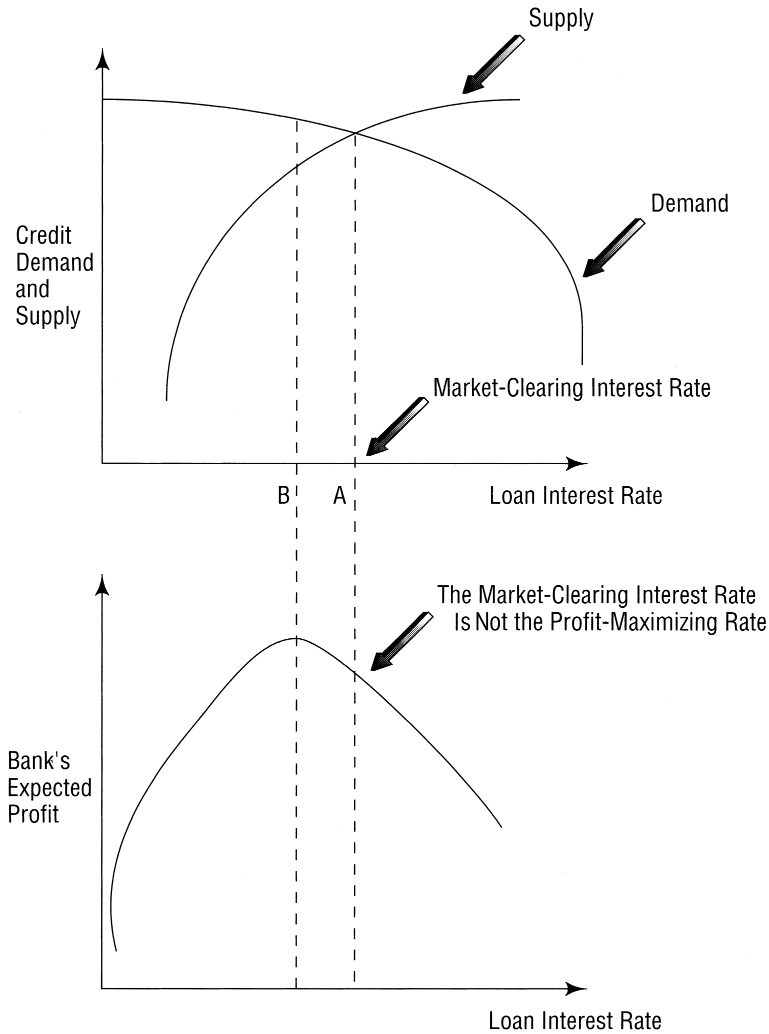


FIGURE 6.6 Credit Rationing

We now provide numerical examples to illustrate these concepts. We will first focus on the *adverse selection* problem, ignoring moral hazard for the moment.

Example 6.2 Suppose that you are the loan officer for the Midtown Community Bank and you know that within a particular risk class, there are two types of borrowers: low-risk borrowers and high-risk borrowers. However, you cannot distinguish between them.

You believe that the probability is 0.5 that a randomly chosen borrower is low risk and 0.5 that the borrower is high risk. There are 1,000 potential loan applications of each type within this risk class. Each applicant would like a loan of \$100. The low-risk borrower will invest this loan in a project that one period hence will yield \$130 with probability 0.9 and nothing with probability 0.1. The high-risk borrower will invest the loan in a project that will yield \$135 with probability 0.8 and nothing with

(Continued)

probability 0.2 one period hence. Midtown Community Bank is a monopolist with respect to these borrowers.[1] Assuming that the only pricing instrument available is the loan interest rate, how should you price a loan to a borrower in this risk class so as to maximize the bank's expected profit? You have only \$100,000 available to lend and the junior lending officer who reports to you has advised you that 2,000 loan applications were received when it was announced that the bank would charge an interest rate of 29 percent. The current riskless rate is 5 percent. Assume that a borrower must have at least 1 dollar of net profit in the successful state in order to apply for a bank loan,[2] and that there is universal risk neutrality.

Solution This example shows how informational considerations can impart rigidity to the bank's loan interest rate. To show this, we proceed in three steps. First, we will compute Midtown Community Bank's expected profit if it charges a rate of interest of 29 percent and is forced to randomly ration half its loan applicants (because all potential borrowers apply). Second, we calculate Midtown's expected profit if it charges a rate higher than 29 percent. In this case, the low-risk borrowers drop out, so that the bank lends only to the high-risk borrowers. Finally, in the third step, we compare the bank's expected profits from the first two steps and show that Midtown Community Bank's expected profit is maximized by setting the loan interest rate at 29 percent and randomly rationing half its credit applicants. The key to this finding is that the bank cannot distinguish between the low- and high-risk borrower.

Step 1 Clearly, if you charge an interest rate of 29 percent, you will have to ration credit since you can lend only \$100,000 to this group of borrowers and the demand is for \$200,000. Now, the *maximum* interest rate that your bank can charge without losing the low-risk borrowers is 29 percent. At this interest rate, the net profit of the low-risk borrower in the successful state is

$$130 - 129 = \$1,$$

because the repayment obligation is \$129. Clearly, the high-risk borrowers will also choose to apply at this interest rate since the net profit of such a borrower in the successful state is

$$135 - 129 = \$6.$$

The total expected profit of Midtown Community Bank, if it lends at an interest rate of 29 percent, is

$$\begin{aligned} &\frac{(0.5 \times 0.9 \times \$129 + 0.5 \times 0.8 \times \$129) \times 1000}{1.05} - \$100{,}000 \qquad [6.3] \\ &= \$4428.57 \end{aligned}$$

The expression in (6.3) can be understood as follows. There is a 0.5 probability that the borrower is low risk, in which case the bank gets repaid \$129 with probability 0.9. Similarly, there is a 0.5 probability that the borrower is high risk, in which case the bank gets repaid \$129 with probability 0.8. This explains the term in the parentheses of the numerator in (6.3). This is multiplied by 1,000 since the bank can make 1,000 such loans. We discount at the riskless rate of 5 percent since the bank is risk neutral. The initial outlay of $100,000 is finally subtracted to arrive at the bank's expected profit.

Step 2 Since there is unsatisfied loan demand at the 29 percent interest rate—half the loan applicants are turned down—it is natural to ask if Midtown can earn a higher expected profit by increasing the loan interest rate.[3]

Clearly, if you raise the loan interest rate above 29 percent, the low-risk borrowers will not wish to borrow. Since only the high-risk borrowers remain, you might as well raise the loan interest rate all the way up to 34 percent, the maximum you can charge the high-risk borrowers before they too drop out. We refer to 34 percent as a *market-clearing* interest rate since at this level, loan demand equals loan supply.[4]

Midtown Community Bank's total expected profit at this interest rate is

$$\frac{0.8 \times \$134 \times 1000}{1.05} - \$100{,}000 \qquad [6.4]$$
$$= \$2095.24.$$

Note that (6.4) recognizes that the bank knows that only the high-risk borrowers will apply.

Step 3 It is clear now that the bank earns a greater profit by charging 29 percent and rationing half its loan applicants rather than raising the loan interest rate to a market clearing 34 percent. This illustrates how adverse selection may cause a profit-maximizing bank to ration credit. Raising interest rates in the face of excess demand may drive away the best customers and leave the bank worse off.

1. We could generalize this example to one in which there are numerous imperfectly competitive banks.
2. This assumption is meant to create a strict incentive for the borrower to apply for a bank loan. In its absence, we could have a situation in which the borrower is indifferent between applying and not applying, and then we would need to assume that an application is made in that case.
3. As the ensuing discussion will make clearer, the loan demand curve in this example is downward sloping in the loan interest rate.
4. Since there are 1,000 high-risk loan applicants and each demands a $100 loan, loan demand will be $100,000.

We now turn to an illustration of the *moral hazard effect*.

Example 6.3 Suppose Midtown Community Bank has received a loan application at $t = 0$ from a firm that currently has no assets except for an investment opportunity available one period hence, at $t = 1$. The customer has stipulated that the loan must be made available at $t = 0$ or not at all. The investment outlay required at $t = 1$ is $I_1 = \$100$, of which $55 will come from a bank loan. The firm will make its decision on whether or not to invest at $t = 1$. The firm currently has some securities outstanding. If the investment is made at $t = 1$, it will yield $\$\tilde{y}$ per year perpetually, beginning at $t = 2$. Although $\tilde{y}$ is not known now, it will be known at $t = 1$. There are five possible states of the world at $t = 1$, as shown in Table 6.3 below.

(*Continued*)

TABLE 6.3 Probability Distribution of $\tilde{y}$

State	Probability	$\tilde{y}$
1	0.05	\$15
2	0.05	\$16
3	0.30	\$17
4	0.40	\$18
5	0.20	\$19

Thus, if state 1 is realized at $t = 1$, the project will pay \$15 per year perpetually beginning $t = 2$.

Assume that the riskless rate is 10 percent and the corporate tax rate is zero. Assuming that \$55 of I_1 will be financed with a loan, and the rest will come from the firm's retained earnings, compute Midtown's expected return as a function of the promised loan interest rate. Assume that I_1 is a perpetual loan (a consol) with interest payable at the end of each period, beginning at the end of the first period, that is, at $t = 2$.

Solution The basic idea conveyed by this example is that it does not benefit the bank to keep increasing the loan interest rate because, beyond some point, an increase discourages the borrower from investing when the bank would prefer to proceed with the project. We solve this problem in three steps. First, we provide a framework for linking the bank's *actual* annual interest payment on the loan as a function of the *promised* interest payment. Second, we calculate the interest payment the bank can *expect* to receive each period for different values of the promised loan interest rate. Finally, in Step 3 we conclude that the bank's expected return is maximized at an "interior" loan interest rate, so that if loan demand exceeds loan supply at this rate, the bank will ration credit rather than raise the loan interest rate further.

Step 1 Since at $t = 1$ all uncertainty is resolved, we can view 10 percent as the appropriate discount rate in determining whether or not to undertake the investment at $t = 1$. That is, I_1 will be made at $t = 1$ if and only if $y_s/0.10 \geq I_1$, where $\tilde{y}_s$ is the share of $\tilde{y}$ accruing to the borrower. If the investment is undertaken, then $\tilde{y}_s = \tilde{y} -$ interest on the \$55 loan. Note that the borrower follows this rule because at the time it has to make the investment (at $t = 1$), it already has the money loaned by the bank, and hence treats it as its own retained earnings.

Let r be the actual annual interest payment on the risky bank loan (viewed at $t = 0$, r is a random variable), assuming a perpetual loan with interest payable every period, beginning at $t = 2$. Let r be the *promised* annual interest payment on any debt outstanding at $t = 0$, where r is promised to begin at $t = 2$.

Note that the bank loan is risky only when viewed at $t = 0$. As mentioned earlier, it becomes riskless at $t = 1$. At $t = 1$ then, the value of the bank's loan is the value of a riskless consol bond with an annual coupon equal to the interest payment the bank knows it will receive perpetually, that is, the value of the bank's loan $= \dfrac{\textit{interest payment}}{0.10}$. For example, at $t = 0$ the promised interest payment to the bank may be \$17, but at $t = 0$ we do not know whether this promise can be kept. But suppose at $t = 1$, state 3 is realized. Then, if the firm adopts the project, the promise can be kept for sure, and the $t = 1$ value of the loan is $\$17/0.10 = \$170$. Alternatively, if state 2 occurs, the promise will not be kept; the bank will receive only \$16 per year perpetually if the project is adopted. Thus, the time 1 value of the loan is $\$16/0.10 = \$160$.

Step 2 Now the expected returns to Midtown with different loan interest payments (choice of investment made at t = 1) are given in Table 6.4 below.

TABLE 6.4 Expected Returns to Bank

Promised loan interest $\tilde{r}$	Minimal level of $\tilde{y}$ for investment I_1, to be made by borrowing firm's shareholders	Probability (at t = 0) that investment I_1 will be made	Expected interest payment on bank loan (view at t = 0)
≤\$5	\$15	1.00	$\tilde{r}$
\$6	16	0.95	\$5.70
\$7	17	0.90	6.30
\$8	18	0.60	4.80
\$9	19	0.20	1.80
\$10	20	0.00	0

In this table, the fourth column is obtained by multiplying each promised payment in the first column by the corresponding probability in the third column. The numbers in the third column are obtained by examining the second column and Table 6.3. The smallest possible $\tilde{y}$ in Table 6.3 is \$15, so that the probability of observing a $\tilde{y}$ greater than or equal to \$15 is 1.00. Similarly, from Table 6.3 we see that the probability of obtaining a $\tilde{y}$ at least as great as \$16 is the probability that the state that will occur is either 2, 3, 4 or 5; this probability is 0.95. The rest of the numbers follow similarly.

Step 3 The above table shows that Midtown Community Bank's expected return *peaks* at a promised loan interest of \$7. Note that the present value of the bank loan at $\tilde{r} = \$7$ is $6.3/0.10 = \$63$, which exceeds the loan amount of \$55; hence, Midtown will be willing to lend. Thus, if loan demand exceeds loan supply at that rate, Midtown will be unwilling to extend more credit even if the borrower offers a higher interest rate. Credit rationing occurs here because of moral hazard. However, this moral hazard is a little different from that discussed earlier, wherein the borrower increased the bank's default risk by switching to a risky project from a safe project. Here the borrower prefers not to invest in a project that would have enhanced the bank's expected return; underinvestment is the problem here.

Bank Capital and Credit Rationing

A bank's capital position also may affect its decision to ration credit since different categories of loans have different capital requirements. Consider a bank that has the necessary deposits but would need to raise additional capital to satisfy a loan request.

The additional cost of raising this capital, relative to that of raising money from other sources, will then be a charge against the bank's profit from making the loan. If this additional cost is sufficiently high, the bank may prefer to invest the available deposits in marketable securities rather than in loans. Many allege that this is what happened in 1990–92 and led to a *credit crunch* in the United States despite monetary policy initiatives aimed at reviving the economy.[10]

We have thus far assumed that the bank and the borrower have a one-period relationship. As pointed out earlier, when the bank and the borrower contract with each other over many time periods, it is sometimes possible to reduce informational problems. Indeed, this is one reason to have long-term bank-borrower relationships.

The Spot Lending Decision

We now turn to the bank's lending decision in light of the possibility of credit rationing. To understand this, we should begin by noting that credit analysis, which is an integral part of the lending decision, is not a binary (0 or 1) process whereby the bank either conducts credit analysis or not. It should more appropriately be viewed as a continuum; the bank can perform credit analysis to varying degrees of detail.

The more elaborate the analysis, the more costly it is for the bank. The point to note is that the degree of elaboration is a matter of choice for the bank and represents an important element of the spot lending decision-making process.

The bank must determine its spot lending policy under uncertainty about both the quantity and the quality of loan demand, and within its own capacity constraints. These constraints include limits on screening and monitoring resources. Consequently, the bank may be unable to accommodate more than a predetermined level of aggregate lending without significantly sacrificing loan quality. Loan quality deterioration may imply an unacceptable elevation in the likelihood of ruin for the bank. This means that the first step in lending policy may be for the bank to establish an upper bound, say $\overline{L}$, on the bank's aggregate lending for a given period, say (0, T).[11] Loan applicants arriving after the bank has reached its loan maximum are presumably rejected indiscriminately, and we refer to this phenomenon as *rationing in the large*. Before reaching its loan maximum, the bank does not ration indiscriminately. Rather, it recognizes applicant attributes and rejects only the less desirable. This phenomenon is referred to as *rationing in the small*.[12] The decision to ration an applicant in the small is predicated on the outcome of the bank's credit analysis and its lending prior to the applicant's arrival, as we shall see below.

Consider now a bank that extends \$1 credit to each randomly arriving customer over a fixed planning period (0, T). If a loan applicant arrives at time t, where $0 \leq t \leq T$, the bank conducts credit analysis to estimate the borrower's repayment

10. Thakor (1996) develops a theoretical model that makes precisely this point, and also provides supporting empirical evidence. The model assumes that the additional cost of capital associated with raising capital is exogenously given, and does not provide an endogenous justification for this cost.

11. In the simplest formulation, this capacity constant, $\overline{L}$, can be thought of as a fixed number of dollars, but a more sophisticated formulation might have this capacity a convex and increasing function of the opportunities the bank perceives.

12. Some refer to "rationing in the large" as a borrower being shut out of the bank credit market entirely and "rationing in the small" as loan rejection by an individual bank. Our usage differs.

probability θ, takes into account cumulative loans made to date, say L_t, and the remaining time until the end of the bank's planning horizon, $T - t$. The bank's spot lending decision can be viewed as an *optimal stopping problem*, that is, the bank must decide when to stop conducting credit analysis and make a decision on whether to grant or deny credit to the applicant based on the available information. Figure 6.7 depicts this decision-making process in a flow chart format.

It is worth noting that at each step, the bank is really making two decisions: (i) whether to acquire and/or process more information about the borrower at additional cost or stop the information acquisition/processing, and (ii) conditional on having decided not to process any more information, whether to extend credit or deny it. Note that these two decisions are made *simultaneously* at each step, rather than sequentially. Moreover, these decisions are affected by L_t and $T - t$. The larger the L_t—the smaller is $\overline{L} - L_t$—the more stringent will be the bank's credit standard (that is, the higher will the estimated θ have to be for the applicant to be granted credit), holding everything else constant. The bank becomes more selective because it has less money to allocate to applicants arriving after t. For similar reasons, the smaller is $T - t$, the more stringent is the bank's credit standard, holding everything else constant. Another important observation is that the *size* of the flow chart (that is the number of steps) in Figure 6.7 is *not* predetermined. Rather, it depends on the information revealed by the credit analysis at each step, as well as L_t and $T - t$. Sometimes, the flow chart will have only one step. Based on a preliminary (and possibly cursory) examination of the borrower, the bank may decide to terminate the

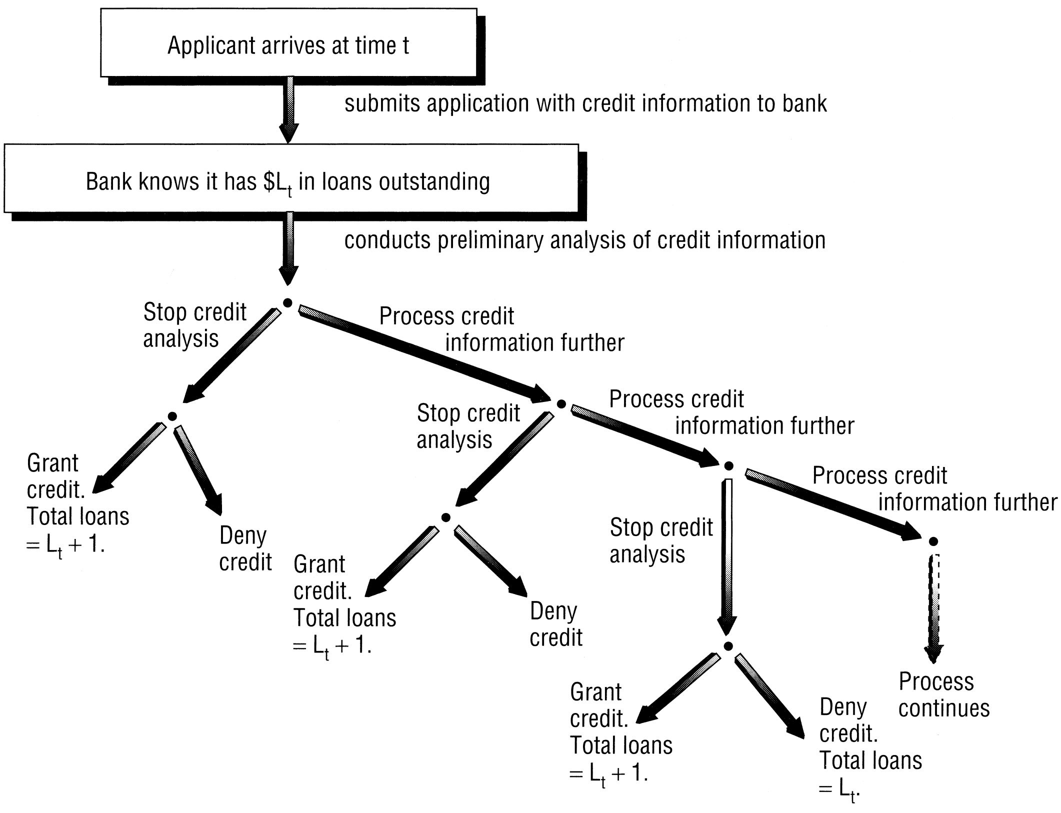


FIGURE 6.7 Flow Chart of the Spot Lending Decision

credit analysis process and either deny credit or grant it. We would expect this to happen in the case of borrowers who are very familiar to the bank either because of their previous credit history or because they belong to some group that contains members with similar default attributes that are relatively well known to the bank. For example, the bank may extend credit to IBM or deny credit to a highly leveraged firm in a risky industry without significant investment in credit analysis in either case. Thus, both intertemporal and cross-sectional reusability of credit information will affect the spot lending decision flow chart. In addition to information about the borrower, L_t and $T - t$ will also affect the size of the flow chart for reasons similar to those mentioned earlier. For example, if $\overline{L} - L_t$ is large and $T - t$ is small, the flow chart may shrink in size as the bank eases its credit standards and grants loans based on favorable results from initial credit analysis.

The amount of information possessed by the bank at the outset about the borrower also has other effects. The bank might charge the borrower a higher interest rate than the breakeven rate that could be charged *given* the bank's information.[13] This is because the bank has better information about the borrower than competing banks do. For example, suppose the information possessed by competing banks indicates that a borrower's default probability is 0.08. Based on its own information, the incumbent bank knows that it is 0.065. Then the incumbent may charge the borrower a rate commensurate with a default-probability of 0.08, thereby earning a positive expected profit due to its informational advantage. We discuss this aspect of bank-customer relationships further in the next section.

Note that the flow chart explains how the bank makes decisions regarding rationing in the small. Once $L_t = \overline{L}$, *all* loan applicants are rationed in the large without any credit analysis.

Some implications of this lending policy perspective are discussed below.[14]

- An increase in $\overline{L}$ will decrease aggregate rationing. This does not mean, however, that each loan applicant will necessarily face a reduced likelihood of rationing. The reason is that the bank will follow a less selective policy from the outset, so that the loans granted by time t will probably be larger. However, it is true that, holding *fixed* L_t, the bank implements a more lax credit standard at time t when a larger $\overline{L}$ is chosen at the outset.
- The effect of $\overline{L}$ on the probability of a *stockout*—the bank exhausts its inventory of loanable funds—at time t is ambiguous. This is because a higher $\overline{L}$ increases lending capacity on the one hand and leads to more lax credit standards on the other. The first effect diminished the stockout probability and the second effect increases it.

Long-Term Bank-Borrower Relationships

In this section we discuss some of the benefits of long-term banking relationships. This will build on our own discussion of relationship lending in Chapter 3. One benefit

13. See Aigner and Sprenkle (1968) for analysis of the bank's optional stopping problem that yields this conclusion.

14. This discussion is based on Deshmukh, Greenbaum, and Kanatas (1983).

is that moral hazard may be reduced. The other is that private information problems can be dealt with more effectively because of information reusability. As we will illustrate in the ensuing discussion, this has potential implications for the design of loan contracts as well as for credit rationing.

Long-Term Relationships and Moral Hazard

When a borrower knows that it may need to borrow in the future, it may limit actions in the current period that would impose losses on the bank. The borrower trades off the current benefits from exploiting the bank against the future costs of poorer credit terms or credit rationing due to these current actions. To see this, consider the following example given in the box below.

Example 6.4 Consider a borrower, Kiddie Toys, Inc., that can choose between two projects, S and R. Project S yields \$150 with probability 0.8 and zero with probability 0.2, whereas project R yields \$162 with probability 0.5 and nothing with probability 0.5. The bank's cost of funds is equal to the riskless interest rate of 5 percent. As a banker, you cannot control your borrower's project choice directly because you cannot observe this choice. You are restricted to making unsecured loans. Assume universal risk neutrality. Moreover, you can charge Kiddie Toys not more than 150 basis points above your breakeven interest rate or it will switch to another bank. Compute the expected payoffs to Kiddie Toys and the bank under the following scenarios: (i) the bank and the borrower can contract with each other over only one period, and (ii) the bank and the borrower can contract over two time periods. In case (i), Kiddie Toys will request a single loan of \$100, and in case (ii), Kiddie Toys will need a sequence of two \$100 loans, with the ability to choose between S and R in each period.

Solution We proceed in four steps. First, we show that in scenario (i) the bank denies credit to Kiddie Toys at any interest rate because it fails to break even regardless of the project chosen by Kiddie Toys. Second, we consider scenario (ii) and show that, by contracting over two periods, it *is* possible for the bank to induce Kiddie Toys to choose S in the second period. For a fixed second-period interest rate that guarantees S will be chosen in the second period, we solve for the maximum interest rate the bank can charge in the first period such that Kiddie Toys will choose S in that period, given that the bank will lend in the second period only if the first-period loan is repaid. Third, given the second-period interest rate in Step 2, we solve for the first-period interest rate needed to permit the bank to break even across its two-period horizon. Finally, in Step 4 we allow the bank to set its first-period interest rate 150 basis points above its breakeven interest rate. We check that Kiddie Toys will choose S in *both* periods and compute the expected profits of the bank and Kiddie Toys.

Step 1 Consider case (i) first. Suppose the bank assumes that Kiddie Toys will choose project R. Then it must set the borrower's repayment obligation at \$105/0.5 = \$210 in order to break even in expected value terms. Given this, Kiddie Toys chooses not to borrow. If the bank assumes that Kiddie Toys will choose S, then it must set its repayment obligation at \$105/0.8 = \$131.25 (an interest rate of 31.25 percent) in

(*Continued*)

order to break even, again in an expected value sense. However, at this interest rate, the expected payoff to Kiddie Toys from choosing S is 0.8(150 − 131.25) = \$15.00, whereas from choosing R it is 0.5(162 − 131.25) = \$15.375. So the bank's belief about the borrower's project choice is contradicted, and it cannot be a Nash equilibrium for the bank to set the loan interest rate at 31.25 percent. Indeed, the maximum interest rate, i_{max}, that the bank can charge such that Kiddie Toys does not strictly prefer R to S is given by the following equation:

$$0.8[150 - (1 + i_{max})100] = 0.5[162 - (1 + i_{max})100].$$

Solving this equation yields $i_{max} = 30$ percent. However, at 30 percent, the bank fails to break even, regardless of the project chosen by Kiddie Toys. Hence, no credit will be extended to the borrower at any interest rate, that is, we have an extreme form of credit rationing. The expected payoff to the bank as well as to the borrower is zero.

Step 2 Now consider scenario (ii). Suppose that as a banker you tell Kiddie Toys: "I'll give you a first-period loan of \$100 at an interest rate of i_1, and a second-period loan of \$100 at an interest rate of i_2, conditional on your repaying the first-period loan. If you default on the first-period loan, then you will not get any second-period credit."

With such a contract, suppose we set $i_2 = 30$ percent. Then we know that the borrower will choose S in the second period. Given this second-period loan interest rate, let i^*_{max} be the maximum value of i_1 such that Kiddie Toys will prefer to invest in S in the first period. Thus, i^*_{max} is the solution to the following equation.

$$\begin{aligned} &0.8\{[150 - (1 + i^*_{max})100] + 0.8 \times [150 - 130]\} \\ &= 0.5\{[162 - (1 + i^*_{max})100] + 0.8 \times [150 - 130]\}. \end{aligned} \qquad [6.5]$$

Note that in (6.5), on the left-hand side we have written Kiddie Toys' expected payoff over two periods from choosing S in the first period, given that S will be chosen in the second period. On the right-hand side, we have written Kiddie Toys' expected payoff over two periods from choosing R in the first period, given that S will be chosen in the second period. In each case we have recognized that second-period credit will be forthcoming only if the first-period project succeeds and the first-period bank loan is repaid; this is done by letting Kiddie Toys' second-period payoff be zero if its first-period project fails and Kiddie Toys consequently defaults on the first-period loan. Solving (6.5) yields $i^*_{max} = 46$ percent.

Step 3 Given a second-period interest rate of 30 percent, let I_1 be the first-period interest rate that the bank needs to charge to break even; remember that at 30 percent, the bank is making an expected loss on the second period loan. Now, $\hat{I}_1$ is the solution to the following equation:

$$[0.8(1 + \hat{i}_1) \times 100 - 105] + 0.8[0.8 \times 130 - 105] = 0 \qquad [6.6]$$

In (6.6), the term $0.8(1 + \hat{I}_1) \times 100 - 105$ is the bank's expected profit on the first-period loan and $0.8 \times 130 - 105$ is its expected profit (which is negative) on the second-period loan. The latter is multiplied with 0.8 (the probability of repayment on the first-period loan) since the second-period loan is made only if the first-period loan is repaid. Solving (6.6) gives $\hat{I}_1 = 32.25$ percent. Note that now the bank is breaking even across two periods rather than in each period.

Step 4 If we assume that on its two-period transaction, the bank can charge 150 basis points above its breakeven rate without losing Kiddie Toys to another bank, then i_1 will be set at 33.75 percent (which is 32.25 percent + 1.5 percent). Kiddie Toys will now choose S in each period. The bank's expected profit over its two-period relationship is given by

$$\begin{aligned}
&0.8(1 + i_1) \times 100 - 105 + 0.8(0.8 \times 130 - 105) \\
&= 0.8 \times 133.75 - 105 + 0.8(0.8 \times 130 - 105) \\
&= \$1.20.
\end{aligned}$$

The expected payoff to Kiddie Toys is given by

$$\begin{aligned}
&0.8(150 - 133.75) + 0.8[0.8(150 - 130)] \\
&= \$25.08.
\end{aligned}$$

As this example illustrates, both the bank and the borrower are better off with a long-term relationship. We go from a situation in which no credit is extended in a single-period relationship to one in which the bank and the borrower negotiate a two-period contract that permits each party to earn a positive expected payoff. The intuition for this improvement is as follows. In the single-period case, it is impossible for the borrower to produce nonnegative expected profit for the bank if it chooses project R, and it is impossible for the bank to induce the borrower to choose project S at an interest rate that permits the bank to break even, assuming that the borrower chooses S. So, no credit is extended. In the two-period case, the bank can commit to a second-period loan at a lower interest rate than it would take for to guarantee that the borrower will choose S in the second period. The bank can recoup this expected loss on the second-period loan by elevating the interest rate on the first-period loan appropriately. This high first-period interest rate will not induce the borrower to choose R in the first period because the borrower is promised a *subsidized* second-period loan only if it repays its first-period loan. This means that the borrower now perceives a *greater cost* to taking risk in the first period than it does in a one-period setting. This creates sufficient room for the desired first-period loan interest rate adjustment by the bank without risking a switch to project R by the borrower.

To recapitulate, a multiperiod relationship with the borrower can mitigate moral hazard.[15] It is less likely that the borrower will exploit the bank when it knows that it must deal with the same bank again. This creates an incentive for bank-borrower relationships.

15. Mitigation of moral hazard through long-term bank-borrower relationships has been examined by Boot and Thakor (1994). See Bhattacharya and Thakor (1993) and Freixas and Rochet (1997) for discussions of the literature.

Three points are worth noting. First, it is important for the bank to offer the borrower a *binding* two-period contract. Since the bank anticipates a loss on its second-period loan, it would prefer not to extend this loan once the second period has arrived. Hence, it is important that a binding contract be negotiated at the outset. Second, as usual, the borrower is free to seek credit elsewhere after the first period. However, no bank will be willing to extend credit to the borrower in a one-period setting, and the incumbent bank is extending a *subsidized* second-period loan. Hence, the borrower will prefer to remain with the same bank for the second period. Finally, it is *time consistent* for the bank to deny the borrower second-period credit, conditional on first-period default, in accordance with the terms of the two-period contract. This is because the bank loses money if it lends in the second period, and will thus do so only if it is bound to do so.

Long-Term Relationships and Private Information

One important advantage of a long-term relationship is that the bank learns about the borrower through time. This lessens the extent to which the borrower is privately informed relative to the bank, and hence improves credit allocations. In other words, the longer a borrower contracts with a bank, the better will be the credit terms it receives. As the borrower keeps repaying the bank, it keeps building an ever-improving track record that enables it to obtain better credit terms through time.[16] We can see this with the following illustration.

Example 6.5 Suppose The Midtown Community Bank is faced with two types of borrowers that it cannot distinguish, G and B. The type-G borrower wishes to borrow \$100 to invest in a single-period project that yields \$135 with probability 0.9 and zero with probability 0.1 at the end of the period. The type-B borrower wishes to borrow the same amount in a project that yields \$150 with probability 0.4 and zero with probability 0.6 at the end of the period.[1] If the borrower comes to the bank for a loan in the second period, it will be to finance exactly the same kind of project as in the first period. Assume that The Midtown Community Bank is perfectly competitive and there is universal risk neutrality. Compute the borrower's interest rates on its first- and second-period loans. Midtown's cost of funds is 5 percent, the riskless rate. Assume that the bank's *prior belief is* that there is a 0.8 probability that the borrower is of type G and a 0.2 probability that it is of type B.

Solution The basic idea is to examine how the bank learns about the borrower through time and how this learning affects the terms of credit. We proceed in four steps. First, we solve for the first-period interest rate that is the same for all borrowers since Midtown cannot distinguish among borrowers. Second, we solve for the break-even second-period interest rate, conditional on first-period project success and loan repayment by the borrower. Repayment of the first-period loan leads Midtown to revise upward its belief that the borrower is of type G. Hence, the second-period

16. See Diamond (1989) and Greenbaum and Venezia (1985).

interest rate in this case is lower than the first-period interest rate. Third, we solve for the breakeven second-period interest rate, conditional on first-period project failure and default. This default leads Midtown to revise downward its belief that the borrower is of type B. This interest rate consequently turns out to be so high that no borrower wishes to take a second-period loan at that rate. Finally, in step 4 we discuss how the first- and second-period rates might actually be determined by Midtown, and the effect of the relative bargaining powers of Midtown and the borrower on this rate.

Step 1 Since The Midtown Community Bank is pooling these two types of borrowers, its breakeven loan interest rate in the first period will reflect the *average* success probability. Let the probability represent the bank's prior belief that the borrower is of type G and let p represent the success probability of a type G borrower. Also let q represent the success probability of a type-B borrower. Then, the average success probability assessed by the bank is given by

$$\gamma p + (1 - \gamma)q = 0.8 \times 0.9 + 0.2 \times 0.4 = 0.8.$$

Hence, the first-period loan interest rate at which the bank breaks even is

$$(1.05/0.8) - 1 = 0.3125 \text{ or } 31.25 \text{ percent.}$$

Step 2 Now, suppose the borrower repays his first-period loan. Then how should Midtown revise its beliefs about the borrower's type? To answer this question, one needs to use *Bayes rule,* which, as we saw in Chapter 1, says that

$$\Pr(x_i|y_i) = \frac{\Pr(y_i|x_i)\Pr(x_i)}{\sum\limits_{i=1}^{n}\Pr(y_i|x_i)\Pr(x_i)} \qquad [6.7]$$

where $x_1, \ldots, x_n$ are the possible realizations of the random variable x and $\Pr(x_i)$ is the prior probability that $x = x_i$, with x_i being some value chosen from $x_1, \ldots, x_n$. Similarly, y_j is some realization of y. In our context, application of Bayes rule means that

$$\begin{aligned}\Pr(\text{borrower is type G}|\text{project succeeds}) &= \Pr(G|S)\\ &= \frac{\Pr(S|G)\Pr(G)}{\Pr(S|G)\Pr(G) + \Pr(S|B)\Pr(B)}\\ &= \frac{p\gamma}{p\gamma + q(1-\gamma)}.\end{aligned} \qquad [6.8]$$

Using (6.8), we see that if there is repayment of the first-period loan, then the bank believes that the probability that the borrower is of type G is given by:

$$\begin{aligned}\Pr(G|S) &= \frac{0.9 \times 0.8}{0.9 \times 0.8 + 0.4 \times 0.2}\\ &= 0.90.\end{aligned}$$

(*Continued*)

Hence, the average second-period success probability is given by:

$$0.9 \times p + 0.1 \times q = 0.9 \times 0.1 + 0.4 = 0.85.$$

The breakeven interest rate of the bank on the second-period loan, conditional on first-period success, is given by $1.05/0.85 - 1 = 23.53$ percent.

Step 3 Note that if there is nonrepayment of the first-period loan due to project failure, then Midtown assesses the probability of the borrower being of type G as (in the equation below, "F" denotes failure)

$$\begin{aligned}\Pr(G|F) &= \frac{\Pr(F|G)\Pr(G)}{\Pr(F|G)\Pr(G) + \Pr(F|B)\Pr(B)} \\ &= \frac{(1-p)\gamma}{(1-p)\gamma + (1-q)(1-\gamma)} \\ &= \frac{0.1 \times 0.8}{0.1 \times 0.8 + 0.6 \times 0.2} \\ &= 0.4.\end{aligned}$$

The bank assesses the *average* success probability for this kind of borrower as

$$0.4 \times p + 0.6 \times q = 0.4 \times 0.9 + 0.6 \times 0.4 = 0.6.$$

Thus, the bank's breakeven interest rate is $(1.05/0.6) - 1 = 75$ percent. But at this rate, neither type would wish to borrow. This means that a borrower who defaults on his first-period loan is effectively denied second-period credit.

Step 4 If the borrower's first-period repayment behavior is freely observable by other banks, then the competitive Midtown Community Bank will charge interest rates of 31.25 percent and 23.53 percent on the first- and second-period loans, respectively. Thus, the loan interest rate declines through time for a borrower who repays his loans. At the other extreme, if competing banks are completely uninformed about the borrower's repayment behavior, then Midtown could charge up to 31.25 percent on the second-period loan and thus make a profit on its second-period loan. Anticipation of this profit could induce Midtown to compete by lowering its first-period loan interest rate below 31.25 percent.[2] Of course, this might strengthen the bargaining power of the borrower who repays his first-period loan. Having paid a lower than breakeven interest on its first-period loan, he knows that the bank need only charge 23.53 percent on the second-period loan to break even on *that* loan. Of course, the borrower had agreed to pay more, but now that promise is "water under the bridge," and (at some cost in terms of his reputation) the borrower could force Midtown to recontract. The interest rate on the second-period loan may end up somewhere between 23.53 percent and 31.25 percent, with the exact interest rate depending on the bargaining strengths of Midtown and the borrower.

1. If the two types of borrowers wished to borrow different amounts and the bank knew which type wanted to borrow how much, the bank would be able to distinguish one type from the other.
2. These issues are analyzed by Sharpe (1990).

In practice, other competing banks do learn something about the borrower, but typically not as much as the incumbent bank. Therefore, through time an *informational surplus* is created in the bank-borrower relationship that could benefit both the incumbent bank and the borrower. Some have argued that this informational surplus could also be socially wasteful.[17] The point is that the incumbent bank's informational advantage could result in its extracting monopoly rents by charging excessively high loan interest rates. This means that the borrower's share of its own project profit is diminished. The borrower's marginal return to working hard to enhance project profits is thereby reduced, and the borrower curtails its effort input. Thus, projects pay off less on average.

Loan Restructuring and Default

We have thus far presented a simplified view of the default process: If the borrower has insufficient cash flow from its project, it defaults. However, as our discussion of bank-borrower relationships has indicated, there is gain from the relationship between the bank and the borrower. Thus, even if we ignore the legal and administrative costs of bankruptcy, the termination of the bank-borrower relationship through default (leading to bankruptcy) is usually costly. The costs to the borrower are transparent. But the bank suffers a cost as well, since a loan default diminishes bank capital. This means that the bank as well as the borrower would be interested in staving off default if possible. This is a major impetus for the widely observed restructuring of bank loans.

There has been extensive research on the issue of default and renegotiation. The basic insights of this research are that the *design* of the debt contract has a lot to do with borrower's incentive to default and the lender's incentive to be willing to renegotiate. Moreover this research has also examined the conditions under which debt contract themselves are the efficient financial contract given the possibility of default and renegotiation.[18]

Types of Financial Distress

Loan restructuring becomes necessary when the borrower is in financial distress. For expositional ease we will classify financial distress into three degrees of severity: mild, moderate, and severe.

(a) Mild Financial Distress: Mild distress is a situation in which the borrower faces the prospect of temporarily insufficient cash flows to service its outstanding debt obligations, but the economic value of the firm comfortably exceeds its repayment obligations. Thus, the borrower faces a temporary cash flow shortfall, rather than insolvency. If forced, the firm could, at some cost, overcome its cash flow deficiency and meet its scheduled debt repayment. Examples are: delaying some investment plans, selling selected assets, or issuing new equity. However, such adjustments could

17. See Rajan (1992).
18. See Hart and Moore (1998). They show that debt contracts are optimal when projects exhibit constant returns to scale and cash flows and asset liquidation value are positively correlated.

diminish the firm's economic value. A less costly alternative may be to approach the lenders with a request to restructure the firm's debt. Lenders, such as banks, may be willing to accommodate such requests for two reasons. First, it signals flexibility on the bank's part and thus improves its reputation in the credit market. Second, to the extent that such an accommodation minimizes borrower value dissipation, the bank may be better off in the long run. Indeed, it can claim for itself a part of the saving achieved by the debt restructuring.

The usual approach to restructuring such loans stretches out the loan's maturity and reduces current interest payments in exchange for an increase in future interest payments. We will discuss two cases of such loan restructurings.

Case 1: Revlon:[19] In 1986, Revlon was acquired by Ronald Perelman, a well-known corporate acquirer, and made a wholly owned subsidiary of Perelman's MacAndrew and Forbes Holdings, Inc. This acquisition was a highly leveraged transaction (HLT), financed with loans from Chemical Bank, Chase Manhattan, Citicorp, and Manufacturers Hanover. An HLT is a loan to a borrower whose debt-to-equity ratio is inordinately high relative to its peers. In particular, it is defined as financing for a buyout, acquisition, or recapitalization that pushes the borrower's liabilities-to-assets ratio to more than 75 percent, or a loan that doubles the company's liabilities and its leverage ratio reaches 50 percent. Revlon had a good record for meeting its financial obligations, and until 1989 it did not appear to be in any danger. However, two events resulted in a mild crisis. First, intense regulatory scrutiny of HLTs in 1990, combined with a deteriorating market for subordinated debt that banks used to augment their capital, caused Revlon's lenders to rethink their position with regard to such loans. The banks decided that they did not want the Revlon loans on their books. They thus designed a refinancing package of four term loans totaling \$1.25 billion and a \$550 million revolving credit facility, and offered these for sale to other lenders. Second, even though Revlon had generally performed well since the Perelman acquisition, many were concerned about its future because of increased competition form Procter & Gamble Company, which had recently acquired Faberge and Elizabeth Arden.

Moody's Investors Service downgraded Revlon's debt rating in January 1990 and noted that industry consolidation "could make maintenance of market shares more difficult and put additional pressure on cash flows." These developments made Revlon's potential creditors nervous.

The refinancing package offered by the original four banks included loans with 4-year maturities, that is, they would come due in 1994. However, \$365 million in Revlon's senior debt would mature in 1995, so banks that bought the refinancing package could find it difficult to help Revlon obtain refinancing in 1994 to repay the 4-year loans. Many potential creditors did not want to deal with a situation in which a subordinated tranche was paid off just before senior lenders were paid. There was additional concern about the refinancing of a \$500 million to \$600 million balloon (principal) payment that would need to be made in 1994.

These difficulties led the four original banks to revise the terms of the deal they were offering to the market. These revisions took the form of structural and pricing adjustments. They were, however, not expected to affect the cost of the loan for Revlon. Rather, any changes in fees or pricing were expected to come out of the pockets of the four banks that underwrote the entire package and would be stuck with any portion of the loan they could not sell.

19. News about Revlon was reported by Lipin (1990a).

This case illustrates some of the difficulties that banks face in restructuring a borrower's debt even when the borrower is in relatively good financial condition. Indeed, Revlon even indicated that asset sales in the next few years were likely and that the resulting cash flows would provide the necessary cushion for complete debt service.

Case 2: Zale Corporation:[20] A Dallas-based jewelry retailer, Zale Corporation, was purchased by Peoples' Jewelers Limited, Toronto, and Swiss-based Swarovski International Holdings AG in late 1986. As in the case of Revlon, the acquisition was financed with considerable debt, making it a HLT. The bank loans used to finance the acquisition were short term. In 1990, Zale was faced with the prospect of repaying these loans. In years past, these loans probably would have been rolled over, with Zale financing its repayment with a high-yield bond issue. However, disarray in the junk-bond market meant that this type of financing was out of the question. Since Zale was not in a position to repay its bank loans without significant impairment to its asset value, it preferred restructuring of its $300 million in acquisition-related debt.

Zale was provided with a restructured $300 million loan commitment maturing in May 1993. This commitment involved unsecured loans, but with the banks being on the same level of seniority as much of the company's high yield from its parents.

Zale illustrates the kinds of steps that borrowers and banks are willing to take to avoid costly default and formal bankruptcy.[21]

(b) Moderate Financial Distress: This is a situation in which default is imminent without debt restructuring. Given the *existing* debt repayment obligations, the economic value of the firm's assets is less than its repayment obligations. However, it is possible that if creditors agree to restructure the debt, the firm could produce sufficient future cash flows so that the *economic value* of the firm's assets would exceed the value of restructured debt, which in turn would exceed the current value of the firm's debt. In this case, the creditor's forbearance is a bet on a change in the company's fortunes. Thus, both the firm's shareholders and its creditors could benefit from the restructuring. The following example illustrates this possibility.

Example 6.6 Marvelous Computers, Inc. currently owes its creditors $120. It is run by an entrepreneur, Mr. Bill Doors, who could manage the firm for one period at a personal cost of $5. Mr. Doors has a unique ability to manage Marvelous Computers; under his stewardship the firm's assets one period from now will be worth $125 with probability 0.9 and $100 with probability 0.1. Under any other management, the firm will be worth $90 for sure, which is its current liquidation value. Assume that the riskless rate is zero and that there is universal risk neutrality. Analyze the possible strategies for the creditors.

Solution There are basically two strategies for the creditors, so that we solve this problem in two steps. First, we analyze what would happen if the creditors insisted on debt repayment on existing terms. Second, we analyze what would happen if the

(*Continued*)

20. News about Zale was reported by Lipin (1990b).

21. Lipin (1990b) quotes Meredith Adler, a high-yield bond analyst at First Boston Corporation, "There's a lot of support from the banks. They like the company and don't want it in bankruptcy."

creditors agree to a restructuring that involves a reduction in Mr. Doors' debt obligation. We find that reducing the face value of the debt increases its economic value to creditors. Hence, restructuring is the preferred strategy.

Step 1 If creditors insist on debt repayment on existing terms, it is clear that Mr. Doors will prefer to default. This is because his payoff conditional on default is zero, whereas if he continues for one more period, his expected payoff is

$$0.9(125 - 120) + 0.1(0) - 5 = -\$0.5,$$

given that the debt obligation must be settled first before Mr. Doors collects anything. Since Mr. Doors' equity in the firm is worth only \$4.50 and the *personal* cost to him of operating the firm is \$5, he computes a payoff of −\$0.50 to managing Marvelous Computers for another period. The creditors' payoff if Marvelous Computers defaults is the liquidation value of the firm, \$90.

Step 2 But now suppose creditors agree to a restructuring whereby the debt repayment obligation of Marvelous Computers is reduced to \$119. Mr. Doors' expected payoff from operating Marvelous Computers for another period is then

$$0.9(125 - 119) + 0.1(0) - 5 = \$0.4,$$

compared to zero in default. Hence, the restructuring provides Mr. Doors with the incentive to continue to operate Marvelous Computers. The value of the debt (the expected payoff to creditors) now becomes

$$0.9 \times 119 + 0.1 \times 100 = \$117.10.$$

Thus, by *reducing* the *face value* of debt by \$1, creditors can *increase* its *economic value* by \$27.10!

We will now see a case of a company in moderate financial distress.

Case 3: The Trump Organization: This company owned and operated a number of hotels (such as the Trump Plaza Hotel) and casinos (such as the Taj Mahal Hotel and Casino), and had over \$2 billion in debt in 1990. On Friday, June 15, 1990, the Trump organization failed to make a \$30 million interest payment to bondholders of Trump's Castle Casino, leaving Mr. Trump ten to thirty days to avoid bankruptcy. Banks, which were major lenders, proposed to postpone some interest payments and provide additional debt financing to enable the Trump organization to avoid bankruptcy.

The four major lenders were the banking units of Citicorp, Chase Manhattan, Bankers Trust, and Manufacturers Hanover. However, there were over 100 additional banks with smaller loans to the Trump organization, and there were also bonds outstanding. The Trump organization's crisis in June 1990, which led to the missed payment, necessitated negotiations between Mr. Trump and the big four banks. Although the banks were nervous about Trump's cash situation, they probably viewed it as prudent not to force Trump property and sell it to repay the notes.

The banks faced a dilemma. On the one hand, they wanted the Trump organization to conserve cash by missing some interest payments on the bank loans as well as on the bonds. On the other hand, they did not want the company to be forced into default by bondholders who could then force a liquidation to collect amounts owed to them. Bondholders had first liens on three of Mr. Trump's properties through first mortgage bonds: the Trump Taj Mahal, Trump Castle Funding, and Trump Plaza Funding.

This was a classic situation in which default seemed imminent without debt restructuring, and yet it seemed to be in the interest of major lenders to forestall default. Indeed, at that time, most of the major lenders seemed confident that their loans to the Trump organization would be sound if default could be avoided.[22] Not surprisingly, the eventual outcome of the negotiations between the Trump organization and its major lenders was that some 80 banks agreed on Tuesday, June 26, 1990, to lend the company an additional $65 million to avoid bankruptcy. The banks also agreed to defer interest payments on $850 million of their $2 billion of outstanding loans.[23]

(c) Severe Financial Distress: This is defined as a situation in which the borrower actually defaults on some debt obligation. A debt restructuring plan may be worked out to preclude formal bankruptcy proceedings. In some cases, the borrower may actually announce its intention to file for reorganization under Chapter 11, and a subset of the lenders may agree to restructure the debt so that a portion of the debt can be repaid and a more efficient reorganization plan can be implemented than one that would be possible if *all* the lenders had to be accommodated. Such a reorganization plan may either be achieved outside of bankruptcy or during bankruptcy proceedings. There are numerous examples of companies that have announced bankruptcies during 2004–05 but continued operating as they reorganized, such as many airlines (e.g., Northwest) as well as companies in the automotive industry (e.g., Delphi). We have already shown that avoiding formal bankruptcy may benefit both the lender and the borrower, but this may not always happen. We will now provide a simple example to show how it may be beneficial for some lenders to help the borrower pay off some of the debt in order to achieve a more efficient reorganization plan.

Example 6.7 Consider Marvelous Computers managed by Mr. Bill Doors. The firm has two kinds of debt outstanding: senior debt under which it owes $100 to bondholders, and a subordinated bank loan that requires a repayment of $1,000. The assets of Marvelous Computers have a current liquidation value of $200, but if the firm continues to operate, it will be worth $1,100 with probability 0.9 and zero with probability 0.1 one period hence. To manage the firm for an additional period, Mr. Doors incurs a personal cost of $5. Mr. Doors has declared that he wishes to

(*Continued*)

22. Lipin and Goodwin (1990) quote an official in the New York office of a major Japanese bank as saying: "We are concerned, but we are still confident with [Mr. Trump's] situation" as far as the developer's ability to make interest payments on his bank debt. They also quote an official with a European bank that was a colender on a $220 million facility for the Trump Palace as saying, "From a financial point of view, I have no problem with the deal."

23. This was reported by Horowitz and Goodwin (1990).

file for bankruptcy and has contacted both the bank and the bondholders' trustee. The bondholders wish to liquidate the firm immediately. What should the bank do? Assume universal risk neutrality and a risk-free interest rate of zero. Mr. Doors owns all of the firm's equity.

Solution We solve this problem in two steps. First, we compute the expected payoffs to all the concerned parties from continuation and liquidation. Second, we examine how the most efficient plan could be implemented. In this example, this is achieved by having the bank buy out the senior debt.

Step 1 It is easy to see why the bondholders prefer immediate liquidation: since the liquidation value of Marvelous Computers is \$200 and they have seniority, they stand to collect \$100, the full amount owed to them. On the other hand, with continuation they receive \$100 with probability 0.9 and nothing with probability 0.1, that is, the expected value of their claim is \$90. From the bank's perspective, however, the expected payoff is $0.9 \times (1100 - 100) = \$900$ if the firm is continued, and \$100 if the firm is liquidated immediately. Mr. Doors also prefers bankruptcy since as a shareholder he collects nothing if Marvelous Computers continues, but the personal cost of continuation is \$5.

Step 2 To ensure that the most efficient investment plan is chosen during bankruptcy, the bank can buy out the senior debt for \$100. Moreover, the bank could agree to restructure the loan so that Mr. Doors owes only \$1,090, instead of \$1,100. Now, the continuation plan will be acceptable to all parties since Mr. Doors' expected payoff is

$$0.9 \times (1100 - 1090) - 5 = \$4,$$

the senior bondholders' payoff is \$100, and the bank's expected payoff is

$$0.9 \times 1090 - 100 = \$881.$$

We will now discuss two cases of severe financial distress.

Case 4: West Point Acquisition Company: This company was the vehicle for Mr. William Farley's acquisition of a number of companies. On March 31, 1990, West Point Acquisition Company defaulted on the payment of \$796 million in principal and interest to a bank group led by Bankers Trust and Wells Fargo & Company. The loan was made to finance the acquisition of West Point-Pepperell, Inc. Earlier, Mr. Farley had obtained a 4-year extension of a separate \$1 billion bridge loan to West Point-Pepperell for operating purposes and this was also due March 31.[24] West Point-Pepperell also had \$900 million in outstanding junk bonds.

24. See Goodwin (1990a). A *bridge loan* is typically made by a commercial or investment bank to provide interim financing for a takeover. A lender must support a bridge loan with capital. It is part of what has come to be known as "merchant banking," which refers to banks taking financial positions in corporate control activity (that is, takeovers and acquisitions).

The banks that loaned West Point Acquisition the money had anticipated the default and had been trying to reach an agreement about how to restructure the loan. It also was reported that the banks wanted to avoid bankruptcy proceedings, but wanted Mr. Farley to reach an agreement with the public holders of the West Point-Pepperell high-yield bonds. Mr. Farley had reportedly offered bondholders a significant equity stake in West Point-Pepperell in exchange for a postponement in interest payments on the debt for up to 3 years.

Bankers Trust and Wells Fargo were also the lead banks on the $1 billion bridge loan, although the composition of the bank group differed from that of the acquisition loan. Apart from the 4-year extension, the bridge loan was restructured with a $165 million *increase* in the amount of credit and a *reduction* in the loan interest rate from prime plus 2.5 percent to prime plus 1.5 percent. This illustrates that lenders may be willing to reduce the actual repayment obligation to increase the *expected* payoff to them.

Case 5: Ames Department Stores, Inc.: On Thursday, April 27, 1990, Ames Department Stores, Inc. announced that it had sought protection from its creditors in federal bankruptcy court by filing for reorganization under Chapter 11 of the Bankruptcy Code.[25] In 1988 Citibank led a bank group that provided $900 million in financing for the purchase of the Zayre department store chain. Hurt by an industry downturn, Ames was in technical default on the $900 million credit agreement and was trying to negotiate a second waiver from the Citibank-led group. Ames said that it filed under Chapter 11 after talks broke down. The basic problem for Ames was apparently the stoppage of shipments to Ames by suppliers who were concerned about the company's cash flow crisis.

At the time of bankruptcy, Ames said that Chemical Bank had agreed to provide it with $250 million of *debtor-in-possession (DIP)* financing. The loan was to be used to repay vendors and fund operations while the company attempted to formulate a reorganization plan. The agreement on DIP financing between Ames and Chemical was, however, subject to court approval. Citibank was also reported to be interested in getting the business. In the box below we provide further details on DIP financing.

Notes on Debtor-in-Possession [DIP] Financing[1]

What exactly are DIP loans, and why have they grown so popular? We discuss these issues here.

Firms filing for bankruptcy often face even greater pressures *after* filing for protection under the bankruptcy laws. These pressures stemmed from suppliers and customers shunning the bankrupt firm because of liquidity concerns. To overcome these difficulties, the 1978 Federal Bankruptcy Code set unified standards for how a debtor could obtain new working capital so that vendors, suppliers, and customers would continue with the company during bankruptcy. The debtor company is protected by freezing both its assets and its liabilities, including working capital bank lines. In place of the corporation, a new legal entity—the debtor-in-possession—is

(*Continued*)

25. See Goodwin (1990b).

created. The 1978 Bankruptcy Code provides incentives for lenders to make *new* debt financing available to the bankrupt firm. It does so by providing a "super priority" lien that gives such a lender a very senior claim on the borrower's cash flow. This claim stands just behind normal administrative expenses but before existing credits, including senior debt. The lien also provides for the loan to mature or be repaid before the debtor emerges from bankruptcy. Some of the key features of DIP loans are as follows:

(1) The DIP lender has claim to any assets not already backing other credits. If assets are insufficient to cover the DIP lender's claim, the DIP lender can make a prior claim on assets already pledged to existing creditors and use them as collateral for the new loan.
(2) Most DIP loans are made as part of loan commitments. Commitment fees range from 2.5 percent to 4 percent of the line and loan interest rates from 1.5 percent to 2.5 percent over prime. In addition, there are usually syndication fees.
(3) Even if the debtor is forced to liquidate while in bankruptcy, the DIP lender is the first to be repaid.

DIP financing is said to have originated in 1984 when Chemical Bank set up a unit to market DIP financing as a new product. The operation began to blossom in 1987 when Texaco, Inc. filed for Chapter 11 protection after losing a $10 billion lawsuit to Pennzoil Company, and turned to Chemical with a $2 billion DIP loan request that was eventually scaled back to $750 million.

Since its inception, the market for DIP lending has become fiercely competitive, but it can also be quite profitable for banks.[2] The United States Supreme Court, in its 2004 decision in *Till v. SCS Credit Corporation*, 1245. Ct. 1951, noted the existence of a free market for lenders advertising financing for Chapter 11 debtors-in-possession. The statutory framework governing DIP loans is Section 364 of Title 11 of the U.S. Bankruptcy Code.

1. See Goodwin (1990b).
2. See Rosenthal (2005) for an extensive discussion.

This case illustrates how lenders may be willing to provide *additional* financing to a borrower unable to repay its existing debt. The reason is as follows. Often a company's cash flow can be impaired by perceptions on the part of its customers, supplier, and possibly creditors that it is in financial distress. In Ames' case, business was disrupted because suppliers stopped shipments. In such cases, it may pay for a bank to either restructure or to infuse additional credit to help the borrower overcome its liquidity shortfall even after the borrower has filed for bankruptcy.

The Coordination Problem in Creditor Coalitions

We have shown how debt restructuring can benefit both the lender and a borrower in financial distress. In most cases, however, the borrower either has borrowed from many lenders or the original lender has sold some pieces of the loan to others. As a result, most debt restructuring plans involve *coalitions* of lenders. This often creates *coordination problems*. It is difficult to ensure that a restructuring plan will be accepted by all creditors, because creditors often have divergent interests. In Example 6.7 we saw how disagreement between two creditors often blocks a restructuring. In that example, it was possible to resolve the conflict by having the junior debt claimant (the bank) buy out the senior debt claimant (the bondholders). However, in practice, efficient resolutions are not always that easy, as the following discussion illustrates.

In the Trump organization case discussed earlier, there were approximately 100 banks involved. Some were "participants"—banks without *direct* relationships with the Trump organization. These banks had purchased loans from the original lenders, referred to as "assignees." When a debt restructuring plan has to be voted on, the assignees cannot vote until they go back and convince the participants. In the Trump case, this persuasion process was protracted and difficult. Many participants apparently asked to be bought out by the assignees. However, the assignees feared that "everyone would want out."[26] And in many cases, "letting a participant out" may be tantamount to providing a free put option. This is illustrated in the following example.

Example 6.8 Having survived earlier travails, Marvelous Computers finds itself in trouble again. It now has three types of debt: a bank loan with the highest priority, senior debt owned by bondholders with the next highest priority, and junior debt owned by bondholders with the lowest priority. The repayment obligations of Marvelous Computers one period hence include the bank loan of \$250, senior bonds of \$45, and junior bonds of \$45. Mr. Doors has announced his intention to declare Marvelous Computers bankrupt. At this stage, creditors must choose one of two mutually exclusive restructuring plans: plan A under which the value of Marvelous Computers next period will be \$290 with probability 0.6 and \$125 with probability 0.4, or plan B under which the value of Marvelous Computers next period will be \$340 with probability 1/3 and \$25 with probability 2/3. If you are the bank's representative, which plan would you prefer and what sort of coordination problems would you expect? Assume universal risk neutrality and a zero discount rate.

Solution We proceed in two steps. First, we calculate the expected payoffs to the various parties from the different plans under the assumption that the absolute priority rule will be strictly observed. Second, we examine the bank's strategies with respect to securing the compliance of junior bondholders to the adoption of the plan preferred by the bank, and discuss the coordination problems that may be encountered.

Step 1 We can readily compute the expected payoffs to the various parties under the assumption that absolute priority rules will be strictly observed. These expected payoffs are given below.

(*Continued*)

26. See Goodwin and Lipin (1990).

TABLE 6.5 Expected Payoffs to Different Claimants

Claimant	Expected Payoff under Plan A	Expected Payoff under Plan B
Bank loan	$200	$100
Senior bonds	$24	$15
Junior bonds	0	$15
Equity (Mr. Doors)	0	0

To understand how these expected payoffs are determined, consider for example the bank's expected payoff under Plan A. With probability 0.6, it is repaid in full ($250) and with probability 0.4, it receives $125; the expected value is $0.6 \times 250 + 0.4 \times 125 = \$200$.

Step 2 Clearly, your bank prefers plan A. Senior bondholders also prefer plan A. However, junior bondholders prefer plan B and will have to be bought out to secure their compliance. Unfortunately for your bank, they *may* insist on being bought out at par rather than at the economic value of their bonds. In this case, your bank and the senior bondholders must pay them $45. In essence, you have given them a free put option with an exercise price of $45! Your bank may find it optimal to pay the $45 since it still leaves you with a *net* expected payoff of $155, which exceeds your expected payoff from plan B. Worse still for your bank, however, senior bondholders may attempt to "free ride" and insist that you buy them out in order to implement plan A. Even though they lose $9 with plan B relative to plan A, they may figure that you have even more to lose with plan B. If your bank buys them out at $45, then they too have been given a free put option. The senior bondholders recognize that even if you buy them out, your net expected payoff with plan A is $110, which exceeds that from plan B.

Renegotiation of Debt Contracts and the Borrower's Choice of Financing Source

We have seen how important renegotiating debt contracts can be to firms in financial difficulty. Moreover, given potential coordination problems in lender coalitions, the degree of renegotiability of debt covenants and other contract features will depend on how many creditors there are and who these creditors happen to be. Debt placed privately with a small number of large investors or a single bank loan may be much easier to renegotiate than public debt. Indeed, widely dispersed debt can significantly raise the costs of renegotiation.[27] This suggests that the borrower should take into account the possibility of future renegotiation of contract terms in choosing its source of credit.[28]

27. See Hart and Moore (1989) for an analysis of optimal debt contracts and renegotiation of the debt contract following default.

28. See Berlin and Mester (1992).

It has been shown that the value of the option to renegotiate debt contracts—the difference in the borrower's net expected profit under a contract when renegotiation is possible and when it is impossible—is high when the firm's ex ante creditworthiness is low.[29] The intuition is that agency problems between shareholders and creditors are likely to be more severe among less creditworthy firms, so that the initial debt covenants to restrict the firm's actions are likely to be relatively restrictive. While restrictive covenants control agency problems, they also reduce the firm's flexibility to pursue profitable investments. Consequently, the importance of renegotiation is elevated for such a firm. This implies that firms with low credit ratings are more likely to negotiate debt contracts with more stringent covenants, but with creditors who are more likely to relax these covenants selectively when they seem inefficient in light of new information. Thus, we would expect firms with poorer credit ratings to take bank loans of privately placed debt and to also accept harsher covenants.

Intermediation Opportunities Created by Financial Distress

One of the reasons why banks might wish to divest loans involving firms in financial distress is that such loans may be classified as risky or nonperforming and thus require more bank capital. Banks may sell these loans to other (possibly nonbank) financial intermediaries that operate under less stringent constraints. An opportunity for financial intermediation is thus created as assets are brokered to those who can hold them more efficiently. It is interesting that this is precisely what has happened as more and more highly leveraged firms have become financially distressed. A mutual fund was established in 1990 with the sole purpose of buying risky loans.[30] Although "vulture funds" that invest in the debt of financially troubled companies have been around for some time, this new mutual fund is the first to purchase *bank loans* exclusively. The fund was initiated by California's Foothill Group and is being marketed to institutional investors by Merrill Lynch & Company. The fund's objective is to purchase both performing and nonperforming loans to distressed and bankrupt companies. These loans are generally senior to high-yield (junk) bonds, which the fund avoids. Although data on the distressed-loans market are not readily available, by May 1990 Foothill had apparently earned 51 percent on the loans. A spokesman for the fund reported at that time that investors could expect to earn 25 percent annually.

Conclusion

This chapter has focused on a variety of issues related to loan pricing, credit rationing, bank-customer relationships, and loan default and restructuring. In an environment in which information "decays" rapidly and new information arrives almost continuously, flexibility is important. Being able to *renegotiate* covenants and other contractual parameters in debt contracts in light of new information becomes essential. Such renegotiation can add value for both the creditor and the borrower.

29. Empirical support is provided by Blackwell and Kidwell (1988).
30. See Lipin (1990c).

Banks have an inherent advantage over capital market financing when it comes to loan workouts and renegotiation of debt contracts. This advantage derives from the bank's position as a "monolithic" lender, whereas capital market financing typically involves many disparate bondholders whose behavior is difficult to coordinate; coordination among creditors is vital to the success of any renegotiation effort. Thus, borrowers who find the option to renegotiate their debt contracts valuable are likely to gravitate to banks for credit. In an intensely competitive environment in which borrower-specific information is volatile, banks would do well to capitalize on their comparative advantage by negotiating restrictive covenants to control agency problems, but also remain flexible enough to accommodate postlending renegotiations of these covenants.

Case Study Zeus Steel, Inc.[31]

Robert Feldon started Zeus Steel, Inc. in December of 1993. He had been a salesman for a large steel fabricator, Seminole Steel Company, prior to forming his own steel fabricating operation. In Mr. Feldon's opinion, Zeus Steel occupies a special position in the local market. Zeus buys "secondary" steel that has been rejected as top grade or "prime" by the steel mills because it is flawed in some way. Because of his long relationship with several suppliers, Mr. Feldon has been very successful in purchasing secondary steel at as much as 33 percent under the going rate for prime steel. Zeus' customers have no objection to using secondary steel either because Zeus removes the flaws (flattens the steel) or because the flaws are only cosmetic (small amounts of rust). The company's primary sources of supply are steel mills, insurance companies (who sell damaged steel that they have insured during ocean shipment), and steel brokers. Often the most difficult time for Zeus is when the steel market is strong and secondary steel becomes very difficult to obtain at a discount. As fabricator, Zeus buys the raw steel and cuts it to order into smaller strips with one of its ten shearing machines.

Feldon started Zeus with \$150,000 of his own money. He purchased a 35-year-old 30,000-square-foot building (with a new overhead crane) for \$60,000 in cash plus \$240,000 to be paid over a 10-year period (\$2,000 per month plus interest at 8 percent); he bought at auction ten used shearing machines for \$100,000, of which he borrowed \$50,000 from the First National Bank (FNB). The remainder of his investment plus a \$50,000 line of credit from FNB was used for working capital.

Robert Feldon, who still owns 100 percent of Zeus Steel, has reached a critical juncture in his relationship with the First National Bank. Phillip Reiling, his old loan officer, has just taken a position at another bank, while his new loan officer, Mike Dickens (MD), has been a commercial loan officer for only 6 months (since his promotion from the credit department). These excerpts from the "credit memoranda" portion of Zeus' credit file reveal the tenuous nature of the banking relationship:

Credit Memoranda

1/30/99 MD
I visited Zeus Steel and met Robert Feldon for the first time. Feldon informed me that he was not at all pleased with his relationship with FNB. According to Feldon, Phillip Reiling had been a good friend but was not always responsive to Zeus' banking requirements.

31. Written by Gregory F. Udell, New York University. We thank Greg for providing us with this case.

Feldon had warned Reiling of Zeus' credit needs many months ago, but nevertheless the $200,000 increase in the line of credit approved last November was treated as a last-minute "crisis." Feldon emphasized that the current $500,000 limit on the line of credit was "strangling" Zeus.

I was given a tour of the plant and was impressed with the level of activity. It seemed as though every square inch of space was being used, much of it to store raw steel. Feldon was quite proud of the fact that he had been able to buy $300,000 of "water logged" coil last month at a bargain rate of $.11 a pound; he apparently already has orders for more than half of that steel.

I told Feldon we'd be more than glad to consider an increase in the Zeus line of credit upon receipt of the 12-31-98 financial statements. Feldon indicated that statements would show an even better year than 1997.

2/26/99 MD
Received urgent phone call from Bob Feldon who indicated that he was about to purchase three new machines for $200,000. He wants FNB to finance the equipment. I suggested lunch on Friday. Feldon agreed to bring an accounts receivable and an accounts payable aging, year-end statements and a new personal statement. Ken Heyden, Bob's accountant, will join us for lunch.

2/28/99 MD
Received a new Dunn & Bradstreet report that revealed some slowness in the trade. Earlier D&B's showed Zeus paying its bill either "discount" or "prompt."

3/2/99 MD
Entertained Bob Feldon for lunch to discuss his request for an increase in the Zeus line and also equipment financing. Also present at the lunch were Ken Heyden and John Garner, head of FNB's Metropolitan Division. Feldon was quite pleased with Zeus' 1998 performance. Much of the increase in sales was due to the acquisition of two new accounts, Archer Manufacturing and Hiawatha Motor Homes. Archer manufactures industrial tool boxes and related accessories that it sells primarily to the construction industry. Hiawatha is in the recreational vehicle business (also a manufacturer). In both cases it was understood that in order to obtain the business, Zeus would have to carry its receivables 60 to 75 days during peak season.

In looking at the statements, we pointed out that it looked as though Zeus was slow in the trade (accounts payable of $1,225,000). Feldon emphasized that with a larger line of credit, Zeus could return to payable its bills in 45 days. Ken Heyden pointed out that his projections indicated that a $750,000 line of credit would be appropriate.

We asked Feldon about the decrease in profit during 1998 and he responded that he just took more out in salary and that his inventory was "understated" for tax purposes. When we expressed concern over the high salary, he said defensively: "You've got my personal guarantee, don't you?"

Feldon reiterated the urgency of his request. The new shearing machines (two 48-inch and one 60-inch) were critical to servicing the two new accounts. We mentioned that we would probably require that the line be secured by accounts receivable and inventory and that FNB normally requires audited financial statements (to which Feldon only half-jokingly responded. "Ken will charge me another $10,000 for that!"). It was appeared that relations are strained.

3/6/99 MD
Contracted three of Zeus' suppliers to check credit. Youngstown and Inland Steel reported that Zeus had been a longtime customer with a good credit experience. Seminole reported that it feels very confident about Feldon but they had experienced slowness up to 60 to 75 days in the Zeus account.

3/7/99 MD
Balance in the Zeus accounts for 1998 were:
Average Collected Balance—$55,000
Average Fee Balance—$17,000

The following meeting took place between John Garner and Dickens on Friday, March 6, 1999, in Garner's office.

Garner: Mike, I'm concerned about Zeus Steel. I know Feldon was irritable and a bit defensive with us last week; but I think he has a right to be. Frankly, this account suffered from neglect under Reiling who took Zeus for granted, keeping Feldon happy with a low interest rate. We might not be able to do everything the way Bob wants, but I believe an honest effort on our part will save the account. After all, there aren't many companies that have grown as dramatically as Zeus. Plus, I've got a lot of respect for Ken Heyden and all the business he's sent our way.

Dickens: A couple of things concern me though. Feldon has taken a lot of money out of Zeus in salary, which has resulted in undercapitalization. With the additional debt he's asking for, I think the ratios will look quite different. I'm also concerned about the company's rapid expansion—I think it may have been at the expense of a sound financial statement.

Garner: We could always bring in a finance company to take the accounts receivable and the inventory as collateral. We could then participate in their line of credit and make the equipment loans ourselves. However, as you know, this is an expensive option for Feldon—the rate on the line will probably jump to 4 percent over prime even with a 50 percent participation on our part. But honestly, I think there are better solutions that are less likely to lose the Zeus business. Zeus has a good profit record and still has a very respectable debt/net worth ratio compared to many of our other local borrowers.

Dickens: We've got to act fast—Feldon needs an answer by Monday and I know he's also talking to Midtown Bank.

Garner: As I see it, our options are: 1) increase the line of credit short of $750,000 on an unsecured basis and approve the equipment loans in accordance with FNB loan policy (75 percent of the purchase price and amortized over three years); 2) approve the full $750,000, but take the A/R and inventory as collateral;[32] 3) approve the equipment loan but get a commercial finance company to do the lien of credit (and buy a participation in that line).

Mike, the choice is yours. You present to the loan committee on Monday morning what you feel is our best offer. If you come up with some other alternative, that's great. All I ask is that you provide the loan committee with a detailed financial analysis in support of your recommendation.

Question: Can you help out Mike Dickens with a financial analysis of Zeus and prepare a recommendation for how the bank should proceed?

32. FNB does not have an asset-based loan department; therefore, if it takes the accounts receivable and inventory as collateral, it must do so without full collateral monitoring.

Financial Statements
(ZEUS STEEL, INC.)
(Prepared without audit by Kenneth Heyden & Company)

Balance Sheet (000's omitted)

Assets		12/31/96		12/31/97		12/31/98
Cash		$ 30		$ 68		$ 24
Accounts Receivable – Net		150		342		698
Inventories (LIFO)		110		326		1006
Other Current Assets		6		8		12
Total Current Assets		296		744		1740
Property, Plant, & Equipment	422		440		490	
Less Accumulated Depreciation	90	332	136	304	188	302
Total Assets		$628		$1048		$2042
Liabilities & Net Worth						
Accounts Payable		$ 60		$ 202		$ 768
Notes Payable – FNB		40		150		500
Current Maturities						
First National Bank		10		10		0
Mortgage		24		24		24
Other Current Liabilities		6		8		26
Total Current Liabilities		140		394		1318
Long-Term Debt						
First National Bank		10		0		0
Mortgage		144		120		96
Total Debt		294		514		1414
Common Stock		150		150		150
Retained Earnings		184		384		478
Total Liabilities & Net Worth		$628		$1048		$2042

Income Statement (000's omitted)

Sales	$1500	$2600	$4300
Cost of Goods Sold			
Beginning Inventory	90	110	326
Purchases	800	1610	3494
Direct Labor	250	274	425
Manufacturing Expenses	54	82	199
Ending Inventory	110	326	1006
Gross Profit	416	850	862
Operating Expenses			
Officer's Salary (Feldon)	100	158	242
Commissions	90	210	290
Office Salaries	30	52	58
Depreciation	42	46	52
Provision for Bad Debts	2	2	24
Miscellaneous	10	16	22
Net Operating Profit	142	366	174
Interest Expense	18	24	38
Net Profit Before Tax	124	342	136
Taxes	38	142	42
Net Profit After Tax	$ 86	$ 200	$ 94

Continued

Projected Income Statement (Zeus Steel, Inc.)

	For the 3 Months Ended			
	3/31/99	6/30/99	9/30/99	12/31/99
Sales	$1400	$1800	$1400	$1400
Gross Profit	350	450	350	350
Operating Expenses	250	320	250	250
Net Operating Profit	100	130	100	100

	Days			
Account Receivable Aging 2/23/99 (Zeus Steel, Inc.)	0–30	31–60	61–90	Over 90
Archer Manufacturing Co.	$79,000	$80,000	$17,000	$
Able Tools Co., Inc.	46,000	52,000		
Centennial Steel Co.	12,000	6,000		
Diversey Products	52,000	38,000	22,000	26,000
Steven's Locker	58,000	48,000		
Hiawatha Motor Homes	76,000	72,000	12,000	
Seminole Steel Co.	42,000	34,000		
Smith Manufacturing Co.	8,000	22,000		
CPN Fabricating	24,000			
Cooper Heating & Cooling	18,000	26,000		
Schiller Manufacturing	30,000	36,000		
Mid-America Products	8,000	10,000	2,000	10,000
Other Accounts (under $10,000)	22,000	54,000	6,000	2,000
Total	$475,000	$478,000	$59,000	$38,000
Total Accounts Receivable: $1,050,000				

Accounts Payable Aging 2/23/99 (Zeus Steel, Inc.)				
Youngstown Steel	$236,000	$72,000	$	$
Seminole Steel Co.	79,000	109,000	40,000	
Inland Steel	101,000	39,000		
Atlantic Underwriters	62,000	107,000	28,000	
Independent Insurance Co.		44,000	30,000	
Robert Cunningham & Co.	57,000	83,000	19,000	
Star Steel	14,000	36,000		
Other Accounts	23,000	27,000	19,000	
Total	$572,000	$517,000	$136,000	$-0-
Total Accounts Payable: $1,225,000				

Personal Financial Statement 2/23/99 (Robert Feldon)

Assets		**Liabilities & Net Worth**	
Cash	$20,000	Notes Payable	$12,000
Marketable Securities (M/V)	270,000	Credit Cards	2,000
Zeus Steel, Inc. (M/V)	2,500,000	Mortgages	
Real Estate (M/V)		Residence	84,000
Residence	300,000	Condominium	75,000
Condominium	220,000		
Personal Property (M/V)	150,000	Net Worth	3,287,000
Total Assets	$3,460,000	Total Liab. & Net Worth	$3,460,000

Continued

LOAN REPORT NUMBER:	1067			**DATE:** 11/19/98
NAME:	Zeus Steel, Inc.			
BUSINESS:	Metal Fabricating			
STARTED:	1993			
PRINCIPALS:	Robert Feldon			
CUSTOMER SINCE:	1993			
OFFICER CONTACT:	PR			
REQUEST:	$500,000 unsecured line of credit (increase from $300,000)			
PURPOSE:	Working capital			
SOURCE OF REPAYMENT:	Collection of Receivables			
DATE:	Prime plus $\frac{1}{2}$% (floating)			
	Compensating balances will be 15% of the line			
AVERAGE BALANCE:				
		1997	1996	1995
	Average Collected	$91,000	$73,000	$46,000
	Average Free	60,000	49,000	31,000
AFFILIATED LOANS:	Auto Loan to R. Feldon –$6,325			
HIGH CREDIT:	$300,000			
PRESENT LIABILITY:	$300,000			
MONTHS OUT OF DEBT (LAST 12 MONTHS)	None			
GUARANTORS:	Robert Feldon (Net Worth $629,000)			
COLLATERAL:	Unsecured			
COMMENTS:				
DATE OF NEXT REVIEW:	3/31/99			

INDUSTRY AVERAGES*

Assets Size	**1 mm–10mm**	**All**
Balance Sheet		
Assets	%	%
Cash & Equivalents	7.2	7.2
Accounts Receivable	25.1	25.9
Inventory	28.1	25.4
Other Current	1.5	1.5
Total Current	61.9	60.0
Fixed Assets (Net)	29.8	31.6
Other Noncurrent	8.3	8.4
Total	100.0	100.0
Liabilities & Net Worth		
Notes Payable Short-Term	8.2	7.1
Current Maturity-L/T Debt	3.4	3.8
Accounts & Notes Payable – Trade	16.1	16.2
Accrued Expenses	6.9	7.7
Other Current	2.6	3.1
Total Current	37.2	37.9
Long-Term Debt	11.7	13.4
All Other Noncurrent	1.5	1.6
Net Worth	49.6	47.1
Total	100.0	100.0

Continued

Income Data

	%	%
Net Sales	100.0	100.0
Cost of Sales	78.6	76.9
Gross Profit	21.4	23.1
Operating Expenses	14.2	16.4
Operating Profits	7.3	6.7
All Other Expenses (Net)	.6	.7
Profit Before Taxes	6.7	6.0

Ratios

Current	1.7	1.7
Quick	.9	.9
Sales/Receivables	9.0	8.9
Cost of Sales/Inventory	6.5	7.2
Cash Flow/Current Maturity	3.8	3.6
Debt/Worth	1.0	1.1
ROE (Before Taxes)	27.6	26.7
ROA (Before Taxes)	12.9	10.8
Sales/Total Assets	2.2	2.2

**Source:* Robert Morris Statement Studies 1998 (Metal Stampings).

ZEUS STEEL, INC.
Financial Analysis

Probability	**1996**	**1997**	**1998**	**Industry (1 mm–10mm)**
Profit	$ 86,000.00	$200,000.00	$ 94,000.00	
Salary	$100,000.00	$158,000.00	$242,000.00	
ROA (Before Taxes)	19.7	32.6	6.6	12.9
ROE (Before Taxes)	37.1	64.0	21.6	27.6
Gross Margin	27.7	32.7	20.0	21.4
Liquidity				
Quick Ratio	1.33	1.06	.56	0.9
Current Ratio	2.11	1.89	1.32	1.7
Turnover				
Accounts Receivable (Days)				
End of Period	36.5	48.0	59.2	41
Average		34.5	44.1	
Inventory				
End of Period	37.0	67.9	106.8	56
Average		45.5	70.7	
Accounts Payable (Days)				
End of Period	27.4	45.8	80.2	
Average		29.7	50.7	
Leverage				
Debt/Worth Ratio	.88	.96	2.25	1.0

Review Questions

1. Suppose a firm has no assets at t = 0, except an option to acquire an investment opportunity at t = 1 for $500 million. The outlay required for this investment will be raised entirely through a bank loan. There are no taxes and everybody is risk neutral. The investment opportunity, if undertaken, will yield a payoff of $X per year perpetually, beginning at t = 2. However, what X will be is *not* known *now*. This knowledge will become available only at t = 1. Right now, we can only describe the possible values of X (at t = 1) by the following probability distribution.

State	Probability	X in millions of dollars
1	0.05	100
2	0.10	150
3	0.15	180
4	0.20	200
5	0.25	210
6	0.25	220

ABC, Inc.

   The riskless rate (single-period) is 10 percent. Draw a graph that shows the relationship between the *current* market value of a perpetual (risky) bank loan for this form and the *promised* interest rate on this loan, which must be paid every year forever, and begins at t = 2.
2. What is credit rationing? Why would it ever be rational for a profit-maximizing bank to ration credit?
3. What are the three main types of financial distress? Why would lenders be willing to restructure debt when the borrower is experiencing mild financial distress? What kinds of accommodations are lenders usually willing to make?
4. What sort of restructuring are lenders willing to engage in when the firm is experiencing moderate financial distress and why?
5. What sort of incentives do lenders have to restructure debt when there is severe financial distress and why?
6. What is a "bridge loan" and how is it related to "merchant banking"?
7. What is DIP financing and why might it be advantageous to existing creditors?
8. Discuss the kinds of coordination problems that can come up in loan workouts and how they might be solved.
9. You are a banker and are confronted with a pool of loan applicants, each of whom can be either low risk or high risk. There are 600 low-risk applicants and 400 high-risk applicants and each applicant is applying for a $100 loan. A low-risk borrower will invest the $100 loan in a project that will yield $150 with probability 0.8 and nothing with probability 0.2 one period hence. A high-risk borrower will invest the $100 loan in a project that will yield $155 with probability 0.7 and nothing with probability 0.3 one period hence. You know that 60 percent of the applicant pool is low risk and 40 percent is high risk, but you cannot tell whether a specific borrower is low risk or high risk. You are a monopolist banker and have $50,000 available to lend. Everybody is risk neutral. The current riskless rate is 8 percent. Each borrower must be allowed

to retain a profit of at least \$5 in the successful state in order to be induced to apply for a bank loan. You have just learned that 1,000 loan applications have been received after you announced a 45 percent loan interest rate. You can satisfy only 500. What should be your optimal (profit-maximizing) loan interest rate? Should it be 45 percent (at which you must ration half the loan applicants) or a higher interest rate at which there is no rationing?

10. Imagine this is January 1, 2002. You are head of the loan department at the High Growth Bank of Los Angeles. Mr. Alex Walker, the founder and CEO of ABC, Inc., a small manufacturing firm, comes to you with a request for a loan that his company will need no later than March 1, 2002. He has indicated that the company will repay the loan February 28, 2003, with principal and interest. ABC's balance sheet and income statement are given below.

ABC, Inc.
Balance Sheet
Year Ended December 31, 2001

Cash	\$50,000
Accounts Receivable	250,000
Due from Mr. Walker	40,000
Inventory	800,000
Total Current Assets	\$1,140,000
Land and Building	\$100,000
Machinery	100,000
Other Fixed Assets	15,000
Total Assets	\$1,355,000
Notes Payable, Bank	\$200,000
Accounts and Notes Payable	300,000
Notes Payable, Assorted Suppliers	100,000
Accruals	50,000
Total Current Liabilities	\$ 650,000
Mortgage	550,000
Common Stock	300,000
Retained Earnings	355,000
Total Liabilities and Equity	\$1,355,000

ABC, Inc.
Income Statement
Year Ended 2001

Net Sales	\$3,650,000
Costs of Goods Sold	2,650,000
Gross Operating Profit	\$1,000,000
General Administrative and Selling Expenses	400,000
Depreciation	20,000
Miscellaneous	200,000
Net Income Before Taxes	\$380,000
Taxes (40%)	152,000
Net Income	\$228,000

In addition to the above information, you have the following ratios, which are averages of the industry to which ABC belongs.

Current ratio	3.0
Inventory turnover ratio	10.0
Average collection ratio	25 days
Fixed-asset turnover ratio	20%
Debt ratio	30%

An important consideration in this loan request is whether or not ABC can internally generate the funds needed to repay the loan by conforming more closely to industry averages. The loan request is for $650,000. You have not determined the loan interest rate yet, but the current annual borrowing rate for this customer is 10 percent. Your expectation is that ABC's borrowing rate over the next few months will stay at about 10 percent. Should you make this loan? If you decide to make the loan, present a qualitative analysis of this loan request and make a summary statement of the necessary loan covenants. There should be at least one affirmative covenant, one negative covenant, and one restrictive clause. You are required to present a brief summary of additional information that could have improved your analysis. (Be specific.)

11. Consider a borrower that can choose between two projects, S and R, each of which will pay off a random amount one period hence. Project S will yield $250 with probability 0.9 and zero with probability 0.1 one period hence. Project R will yield $350 with probability 0.4 and nothing with probability 0.6 one period hence. The bank's cost of funds is equal to the riskless interest rate of 10 percent. As a banker, you cannot control your borrower's project choice directly because you assume universal risk neutrality. Moreover, you can charge this borrower 200 basis points above your breakeven interest rate before the borrower switches to another bank. Compute the expected payoffs of the borrower and the bank under the following two scenarios: (i) the bank and the borrower can contract with each other over only one period and the borrower will request a single loan of $150, and (ii) the borrower will need a sequence of two $150 loans, with the ability to choose between S and R in each period. What should be the choice of the contracting horizon?
12. Consider a firm managed by an entrepreneur. The firm has two kinds of debt outstanding: senior debt under which it owes $150 to bondholders, and a subordinated bank loan that requires a repayment of $1,250. The firm's assets have a current liquidation value of $400, but if the firm continues to operate, it will be worth $1,400 with probability 0.8 and zero with probability 0.2 one period hence. To manage the firm for an additional period, the entrepreneur incurs a personal cost of $25. The entrepreneur has declared that he wishes to file for bankruptcy and has contacted both the bank and the bondholder's trustee. The bondholders wish to liquidate the firm immediately. What should the bank do? Assume universal risk neutrality and a risk-free (discount) rate of zero. The entrepreneur owns all of the firm's equity.
13. Consider a firm that has three types of debt: a bank loan with the highest priority, senior debt owned by bondholders with the next highest priority,

and junior debt owned by bondholders with the lowest priority. The firm's repayment obligations one period hence include the bank loan of \$150, senior bonds of \$60, and junior bonds of \$50. The firm has announced its intention to declare bankruptcy. At this stage, creditors must choose one of two mutually exclusive restructuring plans: plan A under which the value of the firm next period will be \$180 with probability 0.5 and zero with probability 0.5, and plan B under which the value of the firm next period will be \$260 with probability 0.4 and \$20 with probability 0.6. If you are the bank's representative, which plan would you prefer and what sort of coordination problems would you expect? How would you attempt to overcome these problems? Assume universal risk neutrality and a zero discount rate.

14. The following is an excerpt from "A Friendly Conversation." Critique it.

    Appleton: If banks don't do it, someone else will.

    Butterworth: I'm sure that's true, but the question is one of comparative advantage and deadweight losses, that is, reinventing the wheel. For instance, take the example of DIP (Debtor-in-Possession) financing. There's nothing in the law that says only banks can provide it but banks are the biggest players in that market. It's not a mere coincidence.

    Moderator: I guess it's not surprising that the DIP financing market has grown so much, given the debt binge of American corporations in the last decade. I personally find the whole debt restructuring process, and particularly the role of banks in it, quite fascinating. But I do find it ironic that banks are engaged in this at a time when borrowers are complaining about credit rationing by banks.

    Appleton: I think this concern with credit rationing is overdone. First of all, I don't really believe banks ration credit, and if they did, it would be irrational. I'm not in the habit of worrying about why someone may want to smoke a \$5 bill! Moreover, a borrower who is rational could always go elsewhere. But honestly, I have yet to see a convincing study that shows that banks ration credit.

    Moderator: Come now, Alex! Do we need a convincing empirical study substantiating every little truth?

    Butterworth: Please don't answer that, Alex. The fact of the matter is that it *is* possible to explain credit rationing as a rational practice. And this view that a rationed borrower can go "somewhere else" is not surprising coming from you Alex, since you don't believe banks are special anyway.

15. Describe the bank's spot lending process, with particular emphasis on the roles of information-processing-capacity constraints and randomness in loan demand.

References

Aigner, Dennis J., and Case M. Sprenkle, "A Simple Model of Information and Lending Behavior," *Journal of Finance* 23, March 1968, 151–166.

Akerlof, George A., "The Market for 'Lemons": Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics* 84, August 1970, 488–500.

Berlin, Mitchell, and Loretta J. Mester, "Debt Covenants and Renegotiation," *Journal of Financial Intermediation* 2–2, June 1992, 95–133.

Bhattacharya, Sudipto, and Anjan V. Thakor, "Contemporary Banking Theory," *Journal of Financial Intermediation*, October 1993, 2–50.

Blackwell, D.W., and David S. Kidwell, "An Investigation of Cost Differences Between Public Sales and Private Placements of Debt," *Journal of Financial Economics* 22, 1988, 253–278.

Boot, Arnold W. H., and Anjan V. Thakor, "Moral Hazard and Secured Lending in an Infinitely Repeated Credit Market Game," *International Econoner Review* 35–3, November 1994, 899–920.

Deshmukh, Sudhakar, Stuart I. Greenbaum, and George Kanatas, "Lending Policies of Financial Intermediaries Facing Credit and Funding Risk," *Journal of Finance* 38, June 1983, 873–886.

Diamond, Douglas, "Reputation Acquisition in Debt Markets," *Journal of Political Economy* 97, 1989, 828–862.

Freixas, Xavier, and Jean-Charles Rochet, *The Microeconomics of Banking*, MIT Press, October 1997.

Goodwin, William, and Steven Lipin, "4 Main Banks Hold Bag on Trump Loan Payments," *American Banker*, April 3, 1990a.

———, "Chemical to Provide Debtor-in-Possession Financing for Ames Stores," *American Banker*, April 27, 1990b.

———,"4 Main Banks Hold Bag on Trump Loan," *American Banker*, June 14, 1990.

Greenbaum, Stuart I., and Itzhak Venezia, "Partial Exercise of Loan Commitments Under Adaptive Pricing," *Journal of Financial Research* 8, September 1985, 251–263.

Hart, Oliver, and John Moore, "Default and Renegotiation: A Dynamic Model of Debt," *Quarterly Journal Of Economics* 113(1), 1998, 1–41.

———, "Default and Renegotiation: A Dynamic Model of Debt," working paper, MIT, 1989.

Horowitz, Jed, and William Goodwin, "Hanover: Trump Loans Among Nonperformers," *American Banker*, June 28, 1990.

Jaffee, Dwight, and Franco Modigliani, "A Theory and Test of Credit Rationing," *The American Economic Review* 59, December 1969, 850–872.

Lipin, Steven, "Agent Banks Forced to Restructure Revlon Loan," *American Banker*, May 29, 1990a.

———, "Bankruptcy Law Made DIP Loan Appealing," *American Banker*, February 20, 1991.

———, "Fund to Purchase Leveraged Firms' Distressed Loans," *American Banker*, May 17, 1990c.

———, "Junk Bond Ills Force Recasting of Zale Loans," *American Banker*, June 29, 1990b.

Lipin, Steven, and William Goodwin, "Trump's Banks Balked at a Call for Cash, Forcing Negotiations," *American Banker*, June 6, 1990.

Merris, Randall C., "The Prime Rate," *Business Conditions*, The Federal Reserve Bank of Chicago, April 1975, 2–12.

Rajan, Raghuram, "Insiders and Outsiders: The Choice Between Informed and Arm's-Length Debt," *Journal of Finance* 47–2, September 1992, 1367–1400.

Rose, Sanford, "How to Improve Risk-Adjusted Returns," *American Banker*, March 20, 1990.

Rosenthal, Trent L., "Debtor-in-Possession Financing: Opportunities, Risks and Rewards," *The Secured Lender*, May/June 2005, 8–14, 82.

Samuelson, Paul A., Statement in ''Monetary Policy and Management of the Public Debt: Hearings Before the Subcommittee on General Audit Control and Debt Management,'' Joint Committee on the Economic Report, 2nd Congress and Session, 1952.

Sharpe, Steven, ''Asymmetric Information, Bank Lending, and Implicit Contracts: A Stylized Model of Customer Relationships,'' *Journal of Financial Economics* 45–2, September 1990, 1069–1087.

Sprenkle, Case, ''Liability and Asset Uncertainty for Banks,'' *Journal of Banking and Finance* 11, 1987, 147–159.

Stiglitz, Joseph E., and Andrew Weiss, ''Credit Rationing in Markets with Imperfect Information,'' *The American Economic Review* 71, June 1981, 393–410.

Thakor, Anjan V., ''Capital Requirements, Monetary Policy and Aggregate Bank Lending: Theory and Empirical Evidence,'' *Journal of Finance* 51–1, March 1996, 279–324.

Warberg, Carla M., ''Functional Cost Analysis—A New System Approach to Gauging Profitability,'' *Business Review*, Federal Reserve Bank of Dallas, August 1971, 7–11.

CHAPTER • 7

Special Topics in Credit: Syndicated Loans, Loan Sales, and Project Finance

> "The apparently private and technical theme of corporate financing leads us step by step to the heart of major problems of national policy . . . We are dealing here with serious and far-reaching matters which deserve our undivided attention."
>
> *Hans J. Mast, Crédit Suisse*

Glossary of Terms

Syndicated Loan: A loan in which multiple lenders participate.

Project Finance: Financing provided for large projects that are separately incorporated from the sponsoring firm.

Introduction

In the previous two chapters we examined a variety of issues related to bank lending. There are, however, three important topics that we did not cover. These are syndicated lending, loan sales, and project finance. Syndicated lending occurs when multiple lenders participate in making a single large loan. There is a lead lender, typically a commercial bank, in the syndicate that originates the loan and the other lenders participate by providing varying amounts of the loan. A variant of syndicated lending is *loan sales*, which we will also discuss. Project financing occurs when the sponsoring firm for a project decides to incorporate the project as a stand-alone entity outside the firm and seeks financing that has a direct claim on the project cash flows rather than the cash flows of the sponsoring firm. In this chapter we will describe these practices and explain the underlying economic forces at work that make these practices efficient in some circumstances.

Syndicated Lending

In this section we first discuss what a syndicated loan is and the economic functions syndication serves. We then discuss the syndicated loan market, both in domestic and international lending.[1]

What Is Syndicated Lending?

A syndicated loan is a credit granted by a group of lenders, typically banks, to a borrower. Every lender has a *separate* claim on the borrower, even though there is a single loan agreement. There is typically an originating bank (or group of originating banks) that conducts the credit analysis prior to granting the loan and also negotiates the pricing structure of the loan. These originating banks, called the senior syndicate members, are appointed by the borrower and provide the key financial intermediation services of resolving precontract informational asymmetries and designing the loan contract. The others in the syndicate, called the junior banks, provide a portion of the funding. The numbers and identities of the juniors vary depending on the size, complexity and pricing of the loan, as well as the borrower's willingness to expand its banking relationships.

Why do we observe syndicated lending? One of the main reasons is the need for the senior lenders to diversify their credit risk exposure. By inviting banks to participate, the seniors can avoid excessive exposure to a single borrower, while still earning a fee for their origination expertise, including contract design, pricing and distribution services. That is, loan syndication is a way for the bank to solve an inherent tension between the benefits of specialization and the benefits of diversification.

For the junior lenders in the syndicate, syndication enables participation without the costs of origination expertise. That is, these banks can diversify their loan portfolio by adding credits that they lack the expertise to originate themselves. Moreover, it exposes the junior bank to the borrower, and therefore creates the possibility of a future relationship that is deeper and more profitable for the bank.[2]

An example of a syndicated loan structure is provided in Figure 7.1. This syndicated loan took the form of a loan commitment (a topic discussed in greater depth in the next chapter) from a syndicate of banks to Starwood Hotels and Resorts Worldwide, Inc. in 2001. In this syndication, Deutsche Bank AG is the senior bank in the syndicate and Bank One NA, Citibank NA, Credit Lyonnais SA, and UBS AG are the juniors.[3]

1. The discussion in this chapter is based in part on Dennis and Mullineaux (2000) and Gadanecz (2004).
2. See Allen (1990) for a discussion.
3. See Gadanecz (2004).

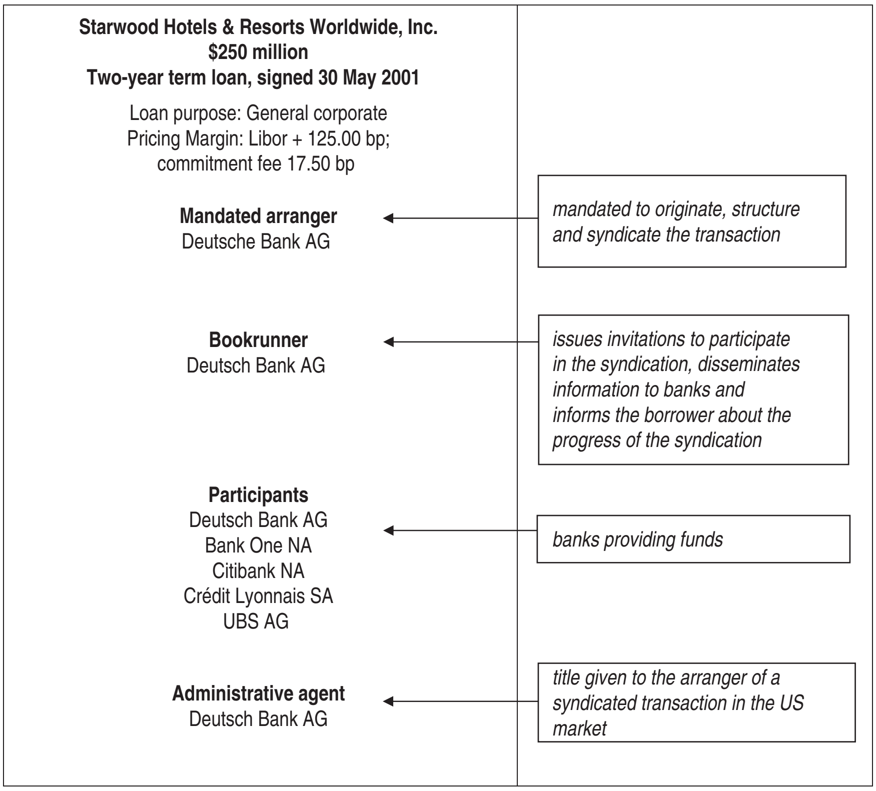


FIGURE 7.1 Example of a Simple Syndicate Structure: Starwood
Source: Dealogic, and Gadanecz (2004).

The Market for Syndicated Loans

Syndicated lending has been very popular in United States domestic lending for many decades. However, since the 1970s, the practice has become an important part of the international lending as well.

In the international market, loan syndications first developed as a sovereign lending business. In fact, just prior to the sovereign default by Mexico in 1982, most of the developing countries' debt consisted of syndicated loans. The repayment difficulties experienced by Mexico and other sovereign borrowers in the 1980s resulted in the restructuring of Mexican debt into Brady bonds in 1989.[4] As a consequence, emerging-market borrowers gravitated toward bond financing, causing a shrinkage in syndicated lending. A revival of syndicated lending occurred in the early 1990s, and now syndicated lending has become the biggest corporate finance market in the United States, as well as the largest source of underwriting revenue for lenders.[5]

Figure 7.2 shows the growth of syndicated lending.

4. A Brady bond is a U.S. dollar-denominated bond issued by an emerging market country, and collateralized by U.S. Treasury zero-coupon bonds. These bonds arose from efforts in the 1980s to reduce the debt burdens of less-developed countries that were prone to default. The bonds were named after U.S. Treasury Secretary Nicholas Brady, who helped international monetary organizations institute the debt-reduction program.

5. See Madan, Sobhani, and Horowitz (1999).

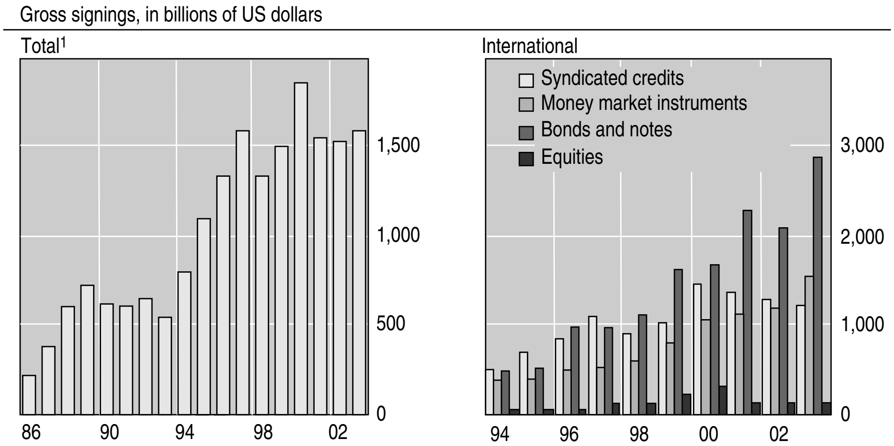


FIGURE 7.2 Syndicated Lending Since the 1980s

[1]Of international and domestic syndicated credit facilities.

Source: Dealogic Loanware; Euromoney; BIS, and Gadanecz (2004).

The Brady plan provided a shot in the arm for the syndicated loan market. By the beginning of the 1990s, banks operating in the syndication loan market had begun applying more sophisticated risk management techniques, making more effective use of covenants and bond-pricing models. A secondary market for loan sales also began to develop as well, which began to attract nonbank financial firms like pension funds and insurance firms. Many banks began to view syndicated lending as a way to gain investment banking business. Moreover, borrowers from emerging markets began to find syndicated loans an attractive alternative and complement to other financing sources. Consequently, syndicated lending grew explosively in the 1990s and new loan signings reached $1.3 trillion in 2003, with borrowers from a wide range of geographies tapping this market. See Figure 7.3.

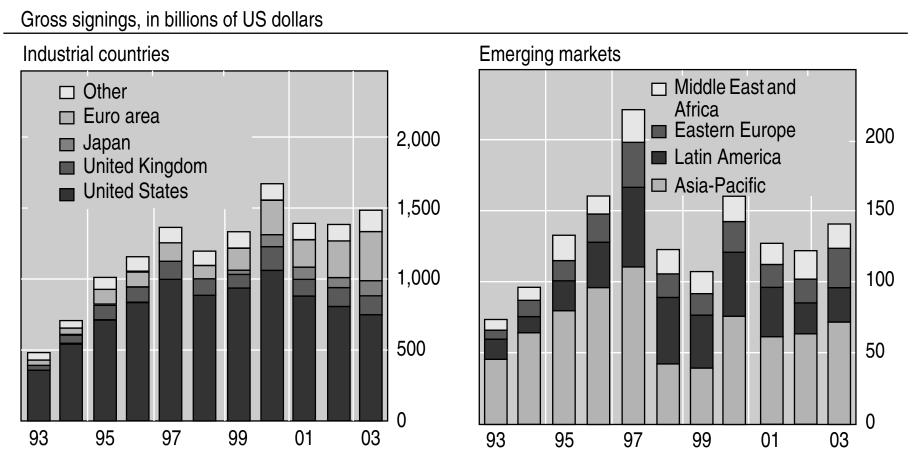


FIGURE 7.3 Syndicated Lending By Nationality of Borrower

Source: Dealogic Loanware, and Gadanecz (2004).

Commercial banks dominate the syndicated lending market, although investment banks became more active in the 1990s. Syndicated loans are also being increasingly traded on secondary markets, facilitated by the standardization of documentation for loan trading and its positive effect on syndicated-loan liquidity.[6]

Participants in the secondary market include: (i) market makers, (ii) active traders, and (iii) occasional sellers/investors. The market-makers are usually large commercial and investment banks. They take positions, commit capital and create liquidity. Active traders are mainly investment and commercial banks, specialized distressed-debt traders and institutional investors, called vulture funds, that trade distressed debt. Less active traders include insurance companies and nonfinancial corporations. Finally, there are also occasional participants who are either buyers or sellers of syndicated loans.

The Brady Plan

The Brady Plan, first announced by U.S. Treasury Secretary Nicholas F. Brady in March 1989, was designed to address the debt crisis of the 1980s that plagued some developing countries. The debt crisis began in 1982, when a number of countries, primarily in Latin America, confronted by high interest rates and low commodities prices, were on the verge of defaulting on their commercial bank loans. This caused credit flows to these countries to dry up, leading to economic stagnation.

The Brady Plan evolved in response to this crisis. Its main features were: (1) bank creditors would grant debt relief in exchange for greater assurance of collectability in the form of principal and interest collateral; (2) debt relief would be linked to some assurance of economic reform and (3) the resulting debt would be easier to trade, to allow creditors greater ability to diversify risk.

Because rescheduling occurred on a case-by-case basis, each Brady issue is unique, but most Brady restructurings included for the lenders a choice between exchanging their loans for either par bonds or discount bonds. Both par and discount bonds were 30-year collateralized bonds. Par bonds represent an exchange of loans for bonds of equal face amount, with a fixed, below-market rate of interest, permitting long-term debt service reduction through concessionary interest terms. Discount bonds represent an exchange of loans for a lesser amount of face value in bonds (generally a 30–50 percent discount), allowing for immediate debt reduction, with a market-based floating rate of interest. The principal of both par and discount bonds was secured at final maturity by a pledge of zero-coupon U.S. Treasury securities denominated in dollars. A portion of the interest payable on par and discount bonds (generally from 12 to 24 months coverage) was also secured by the pledge of high-grade investment securities.

The Brady Plan was successful in many respects. First, it allowed the participating countries to negotiate substantial reductions in their debt service obligations. Second, it helped commercial banks to diversify sovereign risk. Third, it encouraged many developing countries to adopt and pursue ambitious economic reforms. Finally, it has enabled many developing countries to regain access to international capital markets.

6. The professional bodies responsible for initialing such standardization are the Loan Market Association (in Europe) and the Asia Pacific Loan Market Association.

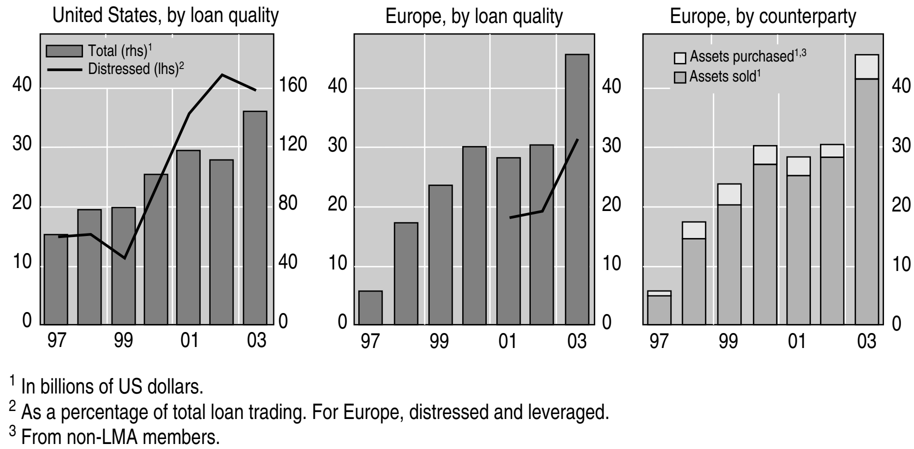


FIGURE 7.4 U.S. and European Secondary Markets for Syndicated Credits

Sources: Loan Market Association (LMA); Loan Pricing Corporation, and Gadanecz (2004).

The biggest secondary market is in the United States, where the trading volume reached $145 billion in 2003, representing 19 percent of originations. Trading volume in Europe was $46 billion. Distressed loans represent a sizeable portion of total secondary trading in the U.S., and are gaining importance in Europe.

Figure 7.4 shows secondary market trading in the U.S. and Europe.

In the Asia-Pacific region, secondary trading volumes are just a small fraction of those in the U.S. and Europe. It is expected, however, that this market will grow.

The Pricing of Syndicated Loans

Syndicated lending is somewhere between a relationship loan (see Chapter 5) and a transaction loan.[7]

The senior bank in the syndicate has a relationship with the borrower and thus, there are aspects of relationship lending that are embedded in a syndicated loan. However, the junior lenders in the syndicate are essentially making transaction loans.

The pricing structure of a syndicated loan resembles that of a loan commitment. A variety of fees are charged, as shown in Table 7.1.

In addition to the fees and the spread between the lending rate and the lender's cost of funds, various mechanisms are used to control risk exposure. These include guarantees, collateral and covenants.

Banks have traditionally sold loans to other banks. Recently, however, their volume has increased dramatically.[8] An increasing number of banks are becoming involved in loan sales as buyers and sellers. Banks commonly employ asset sales specialists. Moreover, the number of banks selling loans through syndication has

7. See Dennis and Mullineaux (2000). Boot and Thakor (2000) distinguish between relationship and transaction lending.

8. See Gorton and Haubrich (1987), Gadanecz (2004), Gorton and Pennacchi (1993) and Pavel and Phillis (1987).

Table 7.1 Structure of Fees in a Syndicated Loan

Fee	Type	Remarks
Arrangement fee	Front-end	Also called *praecipium*. Received and retained by the lead arrangers in return for putting the deal together
Legal fee	Front-end	Remuneration of the legal adviser
Underwriting fee	Front-end	Price of the commitment to obtain financing during the first level of syndication
Participation fee	Front-end	Received by the senior participants
Facility fee	Per annum	Payable to banks in return for providing the facility, whether it is used or not
Commitment fee	Per annum, charged on undrawn part	Paid as long as the facility is not used, to compensate the lender for tying up the capital corresponding to the commitment
Utilization fee	Per annum, charged on drawn part	Boosts the lender's yield; enables the borrower to announce a lower spread to the market than what is actually being paid, as the utilization fee does not always need to be publicized
Agency fee	Per annum	Remuneration of the agent bank's services
Conduit fee	Front-end	Remuneration of the *conduit bank*[1]
Prepayment fee	One-off if prepayment	Penalty for prepayment

[1]The institution through which payments are channelled with a view to avoiding payment of withholding tax. One important consideration for borrowers consenting to their loans being traded on the secondary market is avoiding withholding tax in the country where the acquirer of the loan is domiciled.

Source: Gadanecz (2004) Table 1.

increased, and unlike traditional loan sales, an increasing number of loans (about 60 percent) are now being sold to buyers *outside* the U.S. correspondent banking network, mainly to foreign banks, other intermediaries, and nonfinancial firms. Maturities of loans sold range from one day to two years, with roughly 80 percent having maturities of 90 days or less.

What Is a Loan Sale?

A loan sale is similar to a loan syndication in that the originating bank is able to ensure that part of the funding for the loan comes from other lenders. There are two kinds of commercial loan sales: loan strips and loan participations. A long strip is a short-term share of a long-term loan. When the strip comes due at the end of a given period (say 5, 30 or 60 days), the selling bank must repay the strip holder the contractual amount. In essence, funding has dried up for the loan at that point in time. To continue funding the loan, the origination bank must resell the strip for another period of providing funding itself.

A loan sale without recourse removes the loan from the seller's books and thus does not require reserves or capital to be held against it. The issue is less transparent for strips since they expose the bank to refunding risk. In January 1988, FASB determined that loan strips could be recorded as sales if: (i) the buyer of the strip assumes the full risk of loss, and (ii) the lender has no contractual obligation to repurchase the loan strip.

There is much controversy about whether most loan strips satisfy these conditions. In January 1988, the banking committee of the American Institute of Certified

Public Accounts announced that it would treat a strip as a sale if, at the strip's maturity the original lender can refuse to lend because either: (i) the borrower violates a covenant in the loan contract, or (ii) a material adverse change (MAC) in the borrower's financial condition is discovered. Note that (ii) is the same as the standard MAC clause in loan commitments.

Loan Participation

Like syndicated lending, loan participation is a multilender financing arrangement. It differs from a loan strip in that it is an outright sale of a loan. Participations are loans where the lead lender ("Lead") sells a participation in a loan to one or more participation lenders.[9] The Lead continues to manage the loan on behalf of the participants. The relationship among the lenders is typically formalized in a participation agreement, which stipulates that the participant receives an undivided interest in the loan. The sale of the loan to participants typically occurs after the loan documentation has been executed by the Lead and the borrower. Unlike a syndicated loan, the participants do not contract directly with the borrower. The Lead negotiates the loan terms with the borrower, receives all the payments from the borrower and collateral is maintained by the Lead in its own name. Participants make advances to the Lead, and these take the form of purchases of participation interests.

The advantage of a participant rather than being a junior lender in a syndication is that the lender does not need a separate contract with the borrower and can deal solely with the Lead. Thus, a participation is very much like a pure transaction loan or capital market investment. The advantages of being a junior lender in a syndicate rather than a participant are twofold. One is that the junior lender does not have to worry about the additional risk that the Lead may become insolvent. The other is that the junior lender in a syndicate can hope to develop a relationship with the borrower, something that is less likely for a participant.

From the standpoint of the Lead, one advantage of a loan participation relative to a syndication is that it retains exclusive control over its relationship with the borrower and does not invite potential future competition for relationship lending from the junior lender in the syndicate. The advantage of a syndication for the senior lender is that, because the juniors have direct relationships with the borrowers, the senior lender can free up its own capital in an amount of credit extended by the junior lenders.

Choice Between Loan Syndication and Loan Sales

The syndicated loan market and the market for loan participations have developed because they offer distinct economic advantages for borrowers as well as lenders. For the borrowers, syndicated and participation loans offer some of the advantages of relationship borrowing along with some of the advantages of transaction borrowing (such as liquidity and hence a lower borrowing cost). For the senior lenders, loan syndication permits exploitation of their origination expertise in resolving pre-contract informational asymmetries and negotiating pricing terms, while also enabling them to diversify their credit risk exposure. The same is true for the Lead in a participation loan. For junior lenders, the benefits of loan syndication are the ability

9. See Franks (2005) for discussion.

to diversify into sectors in which they lack origination expertise and to possibly develop a relationship with the borrower that could be deepened in the future. For participants, the benefit of loan participation is the ability to diversify into credits where they lack relationship and/or origination expertise.

Project Finance

In this section we first define project finance, the economic functions it serves, and why it has grown so much recently. We then examine the characteristics of the project financing market.

What Is Project Finance?

Project financing is a technique for financing large-scale infrastructure projects, including those in the natural-resource sectors, like energy and mining. Project financing is different in many respects from conventional financing. With project finance, the firm or public sponsor wishes to invest in a large project, and this is achieved by incorporating the project separately as an independent entity and seeking financing that represents a claim only on the cash flows of the projects. Typically, the sponsor, possibly with other sponsors like investment banks, invests some equity and then finances the rest of the project with debt that is typically *nonrecourse* to the sponsor. Nonrecourse debt means that the lenders have a claim only against the cash flows of the project and not against any other cash flows of the sponsor. The financing mix for the project typically involves a relatively high proportion of debt.

Why is project financing used?[10] There are numerous reasons. First, because the cash flows of the project are *not* commingled with those of the sponsor, it is easier for lenders to resolve the precontract informational asymmetries. This lowers information processing costs for the lender and therefore benefits the borrower. Second, the absence of cash flow commingling also means that asset-substitution moral hazard is reduced. This not only lowers the borrower's cost of capital for financing the project, but also permits higher degrees of leverage to be used, generating a higher debt tax shield. Third, because multiple lenders are involved, the financing structure also has the risk-sharing advantages of syndicated lending. Finally, given the nonrecourse nature of the debt financing for the project, the sponsor does not expose itself to the risk of financial distress in case the project experiences difficulties. This is particulary important for large projects.

There are two reasons why project financing is not used for all projects. First, fixed costs are incurred in establishing a *special-purpose entity* (SPE) to incorporate the project independently. Second, the success of the project typically depends on the joint efforts of many different parties, so there are coordination costs. Project financing is attractive only when its benefits exceed these costs. Although project financing is a venerable practice, it has become an increasingly globalized business since the 1990s.[11] In part this is due to the growing trend to privatize and deregulate many industries around the world.

10. The discussion here is based in part on the theory developed in Shah and Thakor (1987).
11. The following discussion is based in part on Sorge (2004). See also Esty (2003).

The trend in global project financing by geography is shown in Figure 7.5. The significant growth between 1998 and 2000 was in part due to the reallocation of global investors; portfolios from developing to industrialized economies following the East Asian crisis in 1998–99 and new project financing investments in Europe and North America. After 2000, there was a global decline due to general economic slowdown, particularly in the telecommunications and power industries. See Figure 7.6.

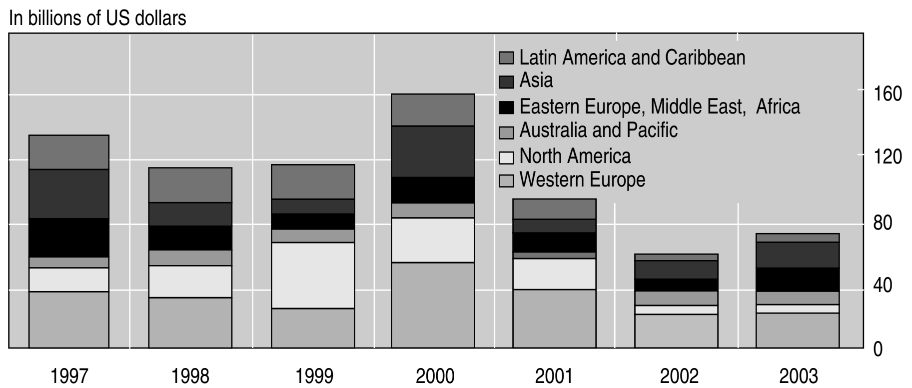


FIGURE 7.5 Project Finance Global Lending By Region

Note: The amounts shown refer to new bak loan commitments for project finance by year and region.
Sources: Dealogic ProjectWare database, and Sorge (2004).

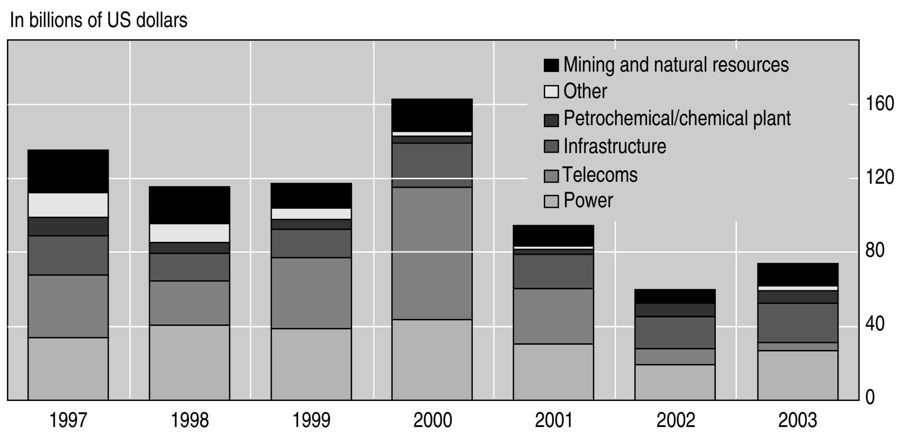


FIGURE 7.6 Project Finance Global Lending By Sector

Note: The amounts shown refer to new bak loan commitments for project finance by year and region.
Sources: Dealogic ProjectWare database, and Sorge (2004).

Despite this downturn, the long-term outlook for project financing is quite bullish. Future demand for infrastructure financing in developing as well as in industrialized countries is apt to grow faster than GDP. It is predicted that during 2005–2025, there will be over 1,500 new electric power generation plants needed in the United States, and developing countries will need annual investment of $120 billion until 2010.[12]

A typical project financing structure is the nexus of multiple contracting relationships as shown in Figure 7.7.

Hybrid structures that combine features of conventional financing and project financing are also being developed. With these structures, the debt financing provided to the project is still nonrecourse to the sponsor, but the idiosyncratic risk of the project is diversified away by lenders who finance *portfolios* of projects rather than single ventures. Moreover, some hybrid structures also involve partnerships between private companies and host governments with private financiers assuming construction and operating risks and host governments taking on market risks.

There are two interesting recent developments in the project finance market. One is the growing popularity of various forms of credit protection such as political risk guarantees, credit derivatives, and a variety of new insurance products that help financiers manage various risks. Second, project finance loans are also increasingly being securitized. This will add considerable liquidity to this market and lower borrowing costs for sponsors.

To summarize, project financing has grown in response to two market forces: (i) the need for borrowers to be able to obtain financing that is exclusively tied to the characteristics of the project and divorced from the sponsor's other cash flows, so as to reduce informational and agency costs, and permit higher leverage; and (ii) the need for lenders to reduce their credit risk exposure by fragmenting the loan for a

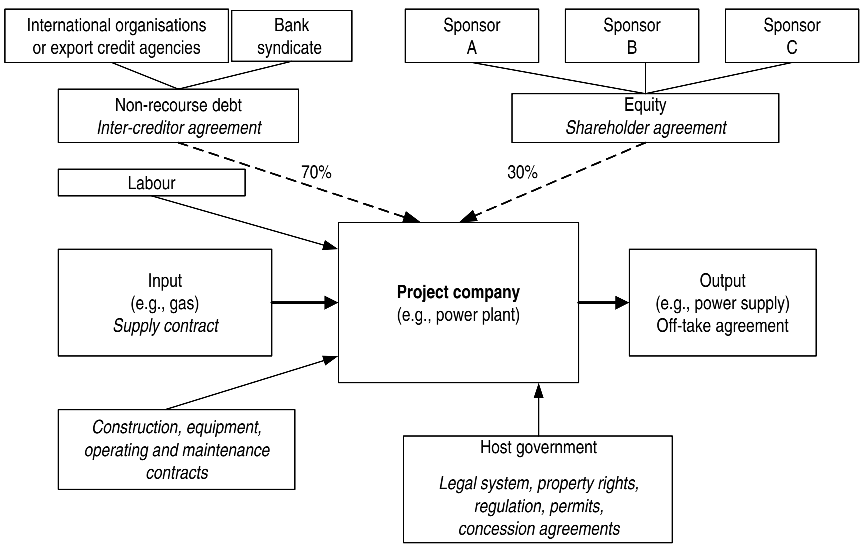


FIGURE 7.7 Typical Project Finance Structure

Note: A typical project company is financed with limited or non-recourse (70%) and sponsors' equity (30%). It buys labour, equipment and other inputs in order to produce a tangible output (energy, infrastructure, etc.). The host goverment provides the legal framework necessary for the project to operate. *Sources*: Adapted from Esty (2003), and Sorge (2004).

12. According to the International Energy Agency.

large project into smaller pieces that are financed by numerous lenders. Project financing is an example of where commercial and investment banks collaborate to provide a variety of brokerage and qualitative asset transformation services, such as resolution of precontract informational asymmetries, reduction of agency costs, and designing the loan contract so as to permit the borrower to obtain more leverage than would be otherwise possible.

Conclusion

In this chapter we have examined two special topics in lending: syndicated lending and project financing. One element that connects them is that project financing usually involves loan syndication as well. Loan syndication creates a loan that combines aspects of a relationship loan and a transaction loan, whereas project financing permits borrowers to undertake large infrastructure projects with significantly higher leverage than would be otherwise possible. Both syndicated lending and project financing involve loan commitments by lenders, a topic we will turn to in the next chapter.

Review Questions

1. What is syndicated lending and what economic functions does it serve?
2. Why is a syndicated loan like a relationship loan and why is it like a transaction loan?
3. What roles do senior and junior lenders play in a syndicated loan?
4. What is project finance and what economic functions does it serve?
5. Why is leverage typically higher in project-financed ventures than in conventional financing?
6. Why is project financing typically used only for very large projects?
7. Why would securitization emerge in project financing? What are the parallels between this and the development of secondary market trading for syndicated loans?

References

Allen, T., 2004, "Developments in the International Syndicated Loan Market in the 1980s," *Quarterly Bulletin*, Bank of England, February 1990.

Boot, Arnoud and Anjan V. Thakor, "Can Relationship Banking Survive Competition?" *Journal of Finance* 55–2, April 2000, 679–714.

Dennis, S. and David Mullineaux, "Syndicated Loans," *Journal of Financial Intermediation*, 9, October 2004, 404–426.

Esty, Benjamin, "The Economic Motivations for Using Project Finance," Working Paper, Harvard Business School, 2003.

Franks, Matthew E., "To Participate or Syndicate? That is the Lender's Question," *The Second Lender*, May/June 2005.

Gadanecz, Biaise, "The Syndicated Loan Market: Structure, Development and Implications," *BIS Quarterly Review*, December 2004, pp. 75–89.

Gorton, Gary and Joe Hanbrich, "Loan Sales, Recourse and Reputation: An Analysis of Secondary Loan Participation," University of Pennsylvania, 1987.
Gorton, Gary and George Pennacchi, "Banks and Loan Sale: Marketing Nonmarketable Assets," manuscript, University of Illinois-University of Pennsylvania, July 1993.
International Energy Agency, "World Energy Investment Outlook," Paris, 2003.
Pavel, Christine and David Phillis, "To Sell or Not to Sell: Loan Sales by Commercial Banks," mimeo, Federal Reserve Bank of Chicago, 1987.
Shah, Sahman and Anjan V. Thakor, "Optimal Capital Structure and Project Financing," *Journal of Economic Theory*, 42, August 1987, pp. 209–243.
Sorge, Marco, "The Nature of Credit Risk in Project Finance," *BIS Quarterly Review*, December 2004, pp. 91–102.

This page intentionally left blank

PART ♦ IV

Off the Bank's Balance Sheet

This page intentionally left blank

CHAPTER ♦ 8

Off-Balance Sheet Banking and Contingent Claims Products

"Has the attention paid to simple capital-asset ratios driven risks off balance sheet, and is off-balance sheet also out of mind?"

Paul Volcker, *Chairman of the Board of Governors of the Federal Reserve system, in an address to the American Bankers Association,* October 1985

Glossary of Terms

Cost of Funds: The effective rate paid by the bank to fund its assets. Source of funds include retail deposits, large-denomination certificates of deposit (CDs), senior and junior debt, preferred stock, and common stock.

Sunk Cost: A cost that has already been incurred and cannot be recovered. Such a cost is irrelevant to a current decision because no matter what the decision, the sunk cost is not affected.

Linear Combination: To simply add up quantities or multiples of quantities. For example, a linear combination of two quantities, say A and B, can be A + B or 3A + 2B, but it is not $A^2 + B$ or $A + \sqrt{B}$.

LIBOR: London Interbank Offer Rate. This is the rate banks charge each other for short-term loans (usually overnight). It is a benchmark interest rate used by banks worldwide.

T-bill Rate: Discount rate on short-maturity debt obligation issued by the U.S. Treasury.

Basis Point: One hundredth of 1 percent.

Basle Accord: An agreement reached among the 12 leading industrialized nations to harmonize international capital standards for banks. It became effective in 1993 and stipulated minimum capital requirements for banks domiciled in all of the signatory nations (see Chapters 2, 11, and 12).

Liability Management: The management of the bank's sources of funding (see Chapter 10).

Derivative: A financial contract, also called a contingent claim, whose value depends on the values of one or more of the underlying assets or indices of asset values. For example, Treasury-bill futures derive their value from movements in the T-bill rate. Bank regulators and banks themselves refer to derivatives more narrowly as contracts, such as forwards, futures, swaps, and options, whose primary purpose is not to borrow and lend but rather to transfer risks associated with fluctuations in asset and liability values.

Initial Public Offering: A public stock offering that converts a privately held firm into a publicly held corporation.

Introduction

Once negligible in amount, and therefore worthy of no more than passing mention in banking texts, *off-balance sheet items* of banks now amount to trillions of dollars in the United States. They include *contingent claims* that represent a variety of exposures across markets and credit risks—standby letters of credit, interest rate and currency swaps, note issuance facilities, options, foreign currencies, fixed- and variable-rate loan commitments, and futures and forward contracts on everything from Treasury bills to gold. Loan commitments are among the largest components of the off-balance sheet items of banks. Also, when added together, off-balance sheet items exceed the total recorded assets of most large banks. This is a little misleading, however, since only some contingent claims impose a (contingent) liability on the bank, and this *contingent liability* is only a fraction of the nominal amount of its outstanding contingent claims. Nonetheless, these data highlight the enormous importance of off-balance sheet items in the current banking environment. The enormous growth in contingent claims of banks has coincided with an explosion in the growth of exchange-traded contingent claims like options and futures. Figure 8.1 *depicts the global growth of exchange-traded options and futures.*

In this chapter we focus on "off-balance sheet (OBS) banking." OBS banking refers to transactions that do not appear on the bank's balance sheet, except possibly as footnotes. OBS items can be divided into two groups: option-like contingent claims and nonoption contingent claims. Table 8.1 shows the various items within each group. Any contingent claim involves a *commitment* on the part of the bank. According to *Webster's* dictionary, a "commitment" is a promise to do something in the future. An option-like contingent claim is a *promise* by the bank to settle in the future at prespecified terms and at the *option* of the holder of the commitment. Thus, an option-like contingent claim imposes a contingent liability on the bank (the seller) and endows the buyer of the commitment with an option. In a competitive market for contingent claims, the bank should be paid a fee at the time the contingent claim is sold that equals the value of the option contained in that claim. Nonoption contingent claims may also involve fees for the bank, but they do not necessarily impose a contingent liability on the bank because there is a symmetry in the obligations of the bank and the customer. Thus, even though there is a future contingency that determines the settlement of the contract, it need not give the customer an option. For example, a forward or futures contract is a nonoption contingent claim.

OBS banking grew explosively in the 1970s and 1980s partly because it was during this period that interest rates and foreign exchange rates became increasingly volatile. This increased volatility in financial and foreign exchange markets and created a strong demand from corporations for financial risk management services. Banks found it profitable to provide these services.[1] Thus, German companies that once borrowed only in D-marks but derived income in other currencies from their foreign operations were now helped by banks to control their foreign exchange risk. Similarly, technology-intensive firms for whom unpredictable short-term revenues imposed severe constraints on research and development (R&D) budgets, approached

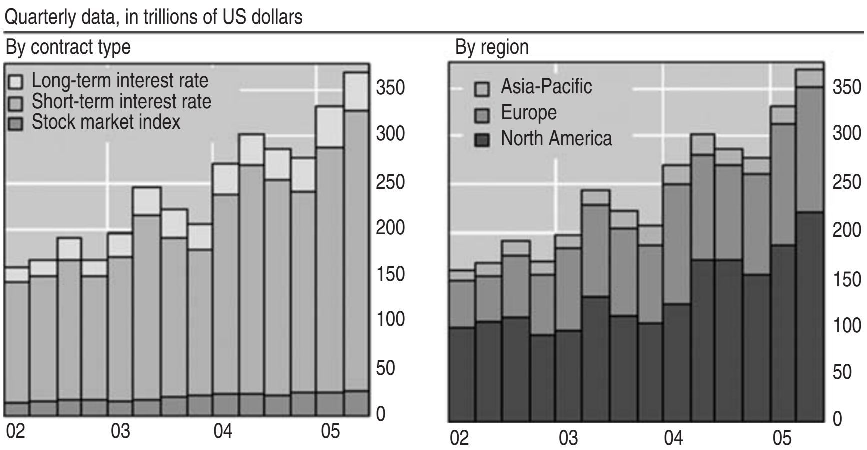


FIGURE 8.1 Volume of Exchange-Traded Futures and Options
Sources: FOW TRADE data; Futures Industry Association; BIS calculations, and Upper (2005).

TABLE 8.1 Off-Balance Sheet Items

Item Classification	Item
Option-like Contingent Claims	1) Loan Commitments and Guarantees
	2) Options
	3) Standby Letters of Credit
Nonoption Contingent Claims	1) Interest Rate Swaps Excluding Those Involving Options
	2) Foreign Currency Transactions Involving Future Settlement
	3) Futures and Forward Contracts

1. *See The Economist* (1993).

banks that provided products designed to hedge overseas income and plan R&D over longer periods. The growth of OBS banking was a natural outgrowth of banks seeking to offer risk management services.

A bank's customer faces two main types of risks. The first is business risk. It may be routine, such as that arising from unpredictable shifts in demand for the firm's output. Or it may be strategic, such as that faced by a defense firm faced with lower demand for arms following the end of the Cold War. The second type of risk is financial. For example, this is the risk of being rationed in the credit market, and the risk of abrupt random movements in interest rates, commodity prices, or currencies. This is where banks enter. They offer loan commitments that can simultaneously guarantee credit availability and interest rate insurance. And banks can offer a variety of derivatives to hedge unpredictable price movements in volatile markets.

While derivatives and other OBS items have been around for a long time,[2] they became widely used only when risk escalated sufficiently. Initially, banks were not involved in the action. Futures and options were offered mainly by organized exchanges such as the Chicago Mercantile Exchange and the Chicago Board of Trade before banks became heavily involved. These were standard contracts for hedging price risk of commodities and later financial claims. However, when corporations wanted products tailored to their specific needs, they turned to banks for those products. This demand led to a variety of custom-tailored contracts such as loan commitments, forward contracts, and swaps.

Banks were interested in custom-designing contingent claims for their clients not only to strengthen customer relationships, but also because sales of contingent claims have proved to be a source of fee income. There are two popularly cited advantages of OBS banking. First, since OBS banking does not involve deposit funding, cash-asset reserves are not needed, and the implicit tax of reserve requirements is avoided. Second, in the past banks were not required to maintain capital against OBS contingencies, although they have been required to do so since the adoption of the guidelines associated with the 1987 *Bank for International Settlements* (BIS) accord.

In the previous three chapters we discussed the spot lending activities of banks. Our focus in this chapter is on *forward* markets. The rest of the chapter is organized as follows. In the next section we describe loan commitments. Economic rationales for the use of loan commitments are provided in the section that follows. Issues related to the valuation (pricing) of loan commitments are examined next. This is followed by a discussion of the differences between exchange-traded put options and loan commitments, and a discussion of the impact of loan commitments on monetary policy. Then, in the next two sections we explain two other contingent claims: letters of credit and interest rate swaps. The issues of risks for banks offering contingent claims are taken up subsequently. The regulatory aspects of contingent claims are taken up next. This is followed by the conclusion of the chapter. A case study is provided to illustrate some of the issues facing a bank that sells contingent claims.

2. Mr. Sykes Wilford, a managing director in Chase Manhattan's risk management group, has been known to show clients a certificate dating from June 1863, when London bankers working for the Confederate States of America raised a dual-currency loan with a coupon linked to future cotton prices.

Loan Commitments: A Description

Definition and Pricing Structure

A loan commitment is a promise to lend up to a prespecified amount to a prespecified customer at prespecified terms. Such a promise is tenable for a prespecified time period (not to be confused with the maturity of the loan). The terms usually specify how the interest rate on the loan will be computed, the maturity of the loan, and the use to which borrowed funds will be put. The bank's compensation for selling the commitment comes in a variety of forms, used in various combinations. It can take the form of a *commitment fee* that is expressed as a percentage of the total commitment and paid up front by the borrower when the commitment is negotiated. It can also take the form of a *usage fee* that is levied on the unused portion of the credit line (for example, 25 to 50 basis points per year). Quite often, commitment and usage fees are employed simultaneously. Also frequently used are *servicing fees* on the borrowed amount to cover the bank's transactions costs, and *compensating balance requirements* that are deposit balances the borrower must keep with the bank during the period of their commitment relationship. These balances are computed as fractions of the total commitment and the bank pays below-market interest rates on these balances.

In Table 8.2 we present data on some actual loan commitments, and Table 8.3 is a detailed description of an actual loan commitment contract. Consider Table 8.2 first. Most striking is the sheer magnitude of some of the individual commitments. To gain perspective, note that the $6 billion commitment to AT&T in 1990 represents about half the total dollar volume of initial public offerings in the United States that year. It is also noteworthy that the margins for banks on many of the loans are quite low. For example, Federal Express was allowed to borrow at the CD rate plus 35 basis points. Since the CD rate is the bank's cost of funds, the bank's margin is only 0.35 percent. In some cases, banks make up for these thin loan margins by charging higher fees on their commitments, but even these fees are slim in many cases. For example, John Fluke Manufacturing paid *no* fees on its commitment and yet had the option to borrow at the CD rate plus 50 basis points. What this table suggests is that the loan commitment market is fairly competitive, at least for larger corporations. Table 8.3 provides details on one of the loan commitments listed in Table 8.2. This contract illustrates a relatively recent innovation in loan commitments, namely offering the customer a choice among rate bases. In this case, Blockbuster Entertainment can borrow at the prime rate, the LIBOR plus 0.5 percent, or the CD rate plus 0.625 percent. The choice increases the customer's flexibility and therefore enhances the commitment's value.

Uses of Loan Commitments

Most business loans are made under loan commitments. These include construction and land development loans, as well as loans to finance leveraged buyouts (LBOs) and mergers and acquisitions (M & A). Loan commitments also include backup lines of credit on commercial paper (the bank agrees to lend to the customer as an alternative to its issuing paper) and note issuance facilities (called NIF, in which the bank agrees to buy the short-term notes of a borrower if the latter is unable to sell them in the markets). Roughly 80 percent of all commercial lending in the United States is done under commitments.

TABLE 8.2 Examples of Actual Loan Commitments

				—Fees (in Basis Points)—					
Commitment Buyer	**Amount**	**Begin**	**End**	**Commitment**	**Annual Servicing**	**Usage**	**Take-down Alternatives**	**Lead Bank**	**Stated Use**
Turner Broadcasting	$200 Million	10-6-89	03-31-92	0	0	62.5	Prime + 75, LIBOR + 175, CD + 187.5	Toronto Division	Commercial Paper Backup
First Brands Corporation	$150 Million	06-15-90	06-15-94	10	0	37.5	Prime, LIBOR + 87.5, CD + 100	Manuf. Hanover	Debt Repayment/ Consolid.
Levi Strauss	$500 Million	03-25-91	06-05-94	12.5	0	0	Prime, LIBOR + 100, CD + 112.5	Bank of America	Debt Repayment/ Consolid.
Safeway Stores	$480 Million	06-12-90	06-20-93	0	0	0	Prime	Banker's Trust	Debt Repayment/ Consolid.
Seagull Energy	$60 Million	04-30-91	07-01-96	0	12.5	17.5	Prime, LIBOR + 87.5, CD + 87.5	Mellon	Debt Repayment/ Consolid.
Action Industries	$30 Million	06-30-88	06-30-91	0	0	37.5	Prime, LIBOR + 87.5, CD + 100	NBD	General Corp. Purposes
Blockbuster Entertainment	$200 Million	08-31-90	09-01-94	0	12.5	12.5	Prime, LIBOR + 50, CD + 62.5	Security Pacific	General Corp. Purposes
J.C. Penney	$750 Million	01-14-91	01-14-94	0	0	18.75	Prime, LIBOR + 37.5	Credit Suisse	General Corp. Purposes
Newmark & Lewis	$30 Million	05-01-90	08-31-90	0	0	0	Prime + 100	Chase Manhattan	General Corp. Purposes
Sharper Image Corporation	$12 Million	05-31-90	06-01-92	0	0	37.5	Prime + 62.5, LIBOR + 200	Continental	General Corp. Purposes

Union Pacific	$550 Million	08-15-88	10-30-93	0	0	15	Prime, LIBOR + 25, CD + 37.5	Citibank	General Corp. Purposes
AT&T	$6 Billion	12-05-90	12-04-91	79.17	13	0	Prime, LIBOR + 37.5, CD + 50	Chemical	Takeover
Federal Express	$150 Million	02-07-90	02-07-94	13	15	0	Prime, LIBOR + 22.5, CD + 35	1st Chicago	Takeover
Ford Motor Company	$800 Million	12-12-89	11-28-92	0	0	0	Prime, LIBOR + 12.5	Citibank	Takeover
Campeau Corporation	$300 Million	12-30-86	12-31-91	243.81	3.05	50	PRIME + 150, LIBOR + 225, CD + 250	Manuf. Hanover	Leveraged Buyout
UAL Corporation	$1.3 Billion	09-13-89	09-13-97	157.64	.69	50	Prime + 100, LIBOR + 200	Citibank	Leveraged Buyout
John Fluke Manufacturing	$37.5 Million	05-04-89	05-04-92	0	0	0	Prime, LIBOR + 50, CD + 50	1st Interstate	Stock Buyback
CUC International	$30 Million	05-25-89	06-01-94	50	0	50	Prime + 150	GE Credit Corp.	Recapitalization
American Oil and Gas	$20 Million	11-13-89	09-30-94	0	0	50	Prime, LIBOR + 300	Prudential	Working Capital
Dunkin Donuts	$35 Million	07-07-89	10-05-89	28.57	0	37.5	Prime, LIBOR + 100	Bank of Boston	Working Capital
L.A. Gear	$150 Million	03-31-89	06-30-90	0	0	50	Prime + 100	Bank of New York	Working Capital
R.H. Macy & Company	$600 Million	04-27-88	04-27-94	150	4.64	50	Prime + 150, LIBOR + 250	Dai-Ichi	Working Capital
Universal Company	$150 Million	06-29-90	03-31-93	0	14.17	0	Prime, LIBOR + 37.5	Sovran	Working Capital

Source: Corporate 10-K and supplemental filings.

TABLE 8.3 Key Terms of a Loan Commitment Contract

Blockbuster Entertainment	
Amount	$200,000,000
Maturity	48 Months
Beginning	8-31-1990
Lender	Security Pacific
Use	General Corporate Purposes
Fee Structure	
Commitment fee	0
Annual servicing fee	12.5 basis points
Usage fee	12.5 basis points
Cancellation fee	0
Take-Down Interest Rate Alternatives	
Prime	
LIBOR + 50 basis points	
CD + 62.5 basis points	

Kinds of Loan Commitments

In addition to use, loan commitments can be classified according to the nature of the interest rate insurance provided to the customer.

Commitments vary in the extent to which they provide interest rate insurance to the borrower. A *fixed-rate* loan commitment gives the customer the right to borrow at an interest rate that is known in advance and hence eliminates all interest rate and availability uncertainty. The more popular *variable-rate* (or fixed formula) loan commitment does not hold the borrowing rate fixed. Rather, it determines the rate according to a formula that involves some index rate. Two common formulas are: additive and multiplicative. The additive version of the variable-rate loan commitment stipulates a borrowing rate that is an index rate at the time of takedown plus a fixed add-on. The less frequently used multiplicative version stipulates a borrowing rate that is an index rate at the time of takedown multiplied by a specified constant. Commonly used index rates are the prime rate, the CD rate, the LIBOR, and the commercial paper rate. Customers may also be offered a choice of formula within a given commitment, for example, prime plus 10 basis points or 1.1 times the CD rate at the time of the borrowing.

Relative to a fixed-rate commitment, a variable-rate commitment does not provide the customer protection against stochastic fluctuations in the index rate. However, as long as there is an element of *fixity* in the borrowing rate, the commitment will have some insurance value to the customer. In the prime-plus commitment, the add-on is held fixed. The customer is thus insured against its add-on being increased due to a possible increase in its credit risk during the commitment period. Likewise, in the prime-times commitment, the multiple is held fixed. In both cases, the customer's commitment borrowing rate at the time of commitment takedown may be lower than the spot rate it would have faced in the absence of the commitment.

Although a loan commitment obliges a bank to lend at a rate below the borrower's spot rate, the bank usually has some latitude in determining whether or not to honor a commitment, even in the case of the most formal agreement. This latitude arises from the adoption of a "general nervous clause" or a "material adverse change" (MAC) clause, which is standard in virtually all loan commitment contracts. This clause allows the bank to dissolve the commitment if the customer's financial condition has "materially" deteriorated between the time the commitment was issued and the time the customer can exercise it. What constitutes material deterioration can, of course, become a legal issue should the denied customer decide to challenge the bank's assessment through litigation. This clause does, however, introduce an element of discretion into the loan commitment contract.[3]

A Summary

We can depict a loan commitment contract as in Figure 8.2. It should be clear by now that a loan commitment is a contingent claim. The contract's contingency hinges upon the interest rate applicable to the specific borrower at the time of commitment takedown. If the spot rate is higher than the commitment rate, the customer will exercise the commitment and the bank will suffer a loss, if only an opportunity loss. If the spot rate is exceeded by the commitment rate, the customer will let the commitment expire unused and borrow instead in the spot market.[4] Thus, the bank has an obligation and the customer has an option. The bank has a loss in

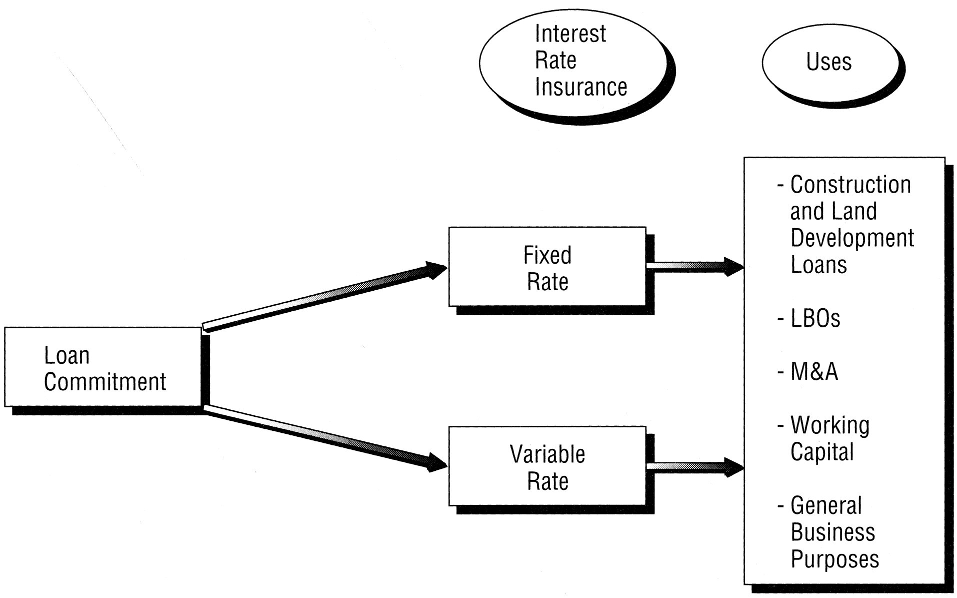


FIGURE 8.2 Depiction of Loan Commitment Classification

3. Boot, Greenbaum, and Thakor (1993) explore the reasons for this discretion and Boot, Thakor, and Udell (1991) examine the manner in which that discretion affects the loan commitment market. See also Holmstrom and Tirole (1993).

4. A usage fee alters this simple decision rule. In its presence, the customer will access the spot market only if the *effective* cost of spot borrowing—and that includes the price the customer must pay for not using the line—is lower than the commitment borrowing rate.

those states of nature in which the customer will exercise the commitment, and this loss is contingent on the occurrence of those states of nature.

Rationale for Loan Commitments

In this section we offer several explanations for the growing importance of commitments.

Supply-Side Explanations

The supply-side explanations for loan commitments attempt to shed light on the popularity of loan commitments by examining the incentives that banks (the suppliers of loan commitments) have to sell the contracts.

(a) Regulatory Taxes: Some believe that loan commitments have been popular because they permit banks to generate fee revenue without having to keep additional capital to support the loan commitments. Moreover, until the commitment is actually taken down, there is no loan, which means no funding has actually taken place. Consequently, the bank needs no deposits until commitment takedown, which implies that reserve requirements[5] do not affect the commitment until that time. In fact, if the bank is interested only in generating the fee income related to the loan commitment, it could sell the commitment and avoid funding the (potential) loan under the commitment altogether. This could be achieved by selling the loan to another bank if and when the customer decides to exercise the commitment. Similarly, the bank could securitize the loan.

(b) Contractual Discretion and Reputation: Another supply-side explanation for the growth in contingent claims relies upon the notion that banks face a trade-off between financial and reputational capital. Simply put, it says that since contingent claims are promises to deliver something in the future, but invariably involve "escape clauses" that introduce contractual discretion and permit the bank to not honor its promises under "extenuating" circumstances, issuing such claims gives the bank improved ability to manage its overall portfolio of *financial* and reputational capital.[6] Consider a bank that has built up a reputation for honoring its contingent claims even in circumstances where provisions in the terms of its contract with the other party would give it the latitude not to. For example, a bank may have agreed to a $100 million credit line at 10 percent interest to a customer whose spot borrowing rate at the time of commitment takedown is 15 percent and whose financial condition at that time is sufficiently murky to enable the bank to invoke the MAC clause and deny credit. Yet a bank with sufficient financial capital may permit the customer to exercise the commitment because this allows the bank to build up its reputational capital. Such reputational capital is of value since it enables the bank to sell *future* contingent claims at higher prices. Now suppose a bank that has accumulated quite a bit of such

5. Banks are required to hold "reserves" against their deposit liabilities (see Chapter 2). Various assets qualify as legal reserves, including cash in vault, deposits with the Federal Reserve, and so on. These reserve requirements vary depending on the nature of the bank's deposit liability. We will have more to say about this in later chapters. However, see Kareken (1987) for a lucid account of the inadequacy of regulatory taxes like reserve requirements to explain the growth in contingent claims.

6. This explanation is based on the theory in Boot, Greenbaum, and Thakor (1993).

reputational capital but finds financial capital in scarce supply is faced with the same decision. Such a bank may well decide to invoke the MAC clause and not honor the commitment. This will result in some depreciation of its reputational capital, but it will conserve scarce financial capital. Thus, the decision to not honor the commitment can be seen as an optimal trade-off by the bank between its reputational and financial capital, and it is essentially an act of *liquefying* its reputational capital; note that, unlike the bank's financial capital, its reputational capital cannot (with good reason) be *directly* traded. In a later section, we discuss a Security Pacific interest rate swap deal where contractual discretion is employed to write down reputational capital.

You should note that the ability to introduce discretion into a contract is predicated on the contract involving the promise of *future* delivery, as a contingent claim does. Moreover, discretion in the loan commitment contract is beneficial because it permits the bank to trade off liquid against illiquid assets.

(c) Demand Forecasting: By participating in the loan commitment market, the bank can obtain valuable information about future loan demand.[7] The reason is that customers will purchase commitments for amounts that are related to their expected future borrowing needs. This permits the bank to plan its funding and other activities accordingly. The question we turn to next is why customers would demand loan commitments.

Demand-Side Explanations

Demand-side explanations focus on the benefits of loan commitments to the purchaser. Many benefits have been identified, five of which are discussed below.

(a) Risk-Sharing Considerations:[8] As discussed in Chapter 4, banks sometimes mismatch their balance sheets in order to profit from the term premiums in the term structure of interest rates. That is but one way for banks to increase expected profits by taking on interest rate risk. Loan commitments provide another. When a bank sells a fixed-rate loan commitment, it accepts the interest rate risk that the customer would otherwise bear if it were to borrow in the spot market for credit. The customer should, of course, be willing to compensate the bank for taking this risk, and this compensation should be reflected in the price paid for the loan commitment.

Borrowers who are more risk averse than the bank should be willing to pay the bank for taking interest rate risk on their behalf. In other words, the risk premium demanded by the bank for bearing interest rate risk will be lower than that demanded by the customer for bearing the same risk if the latter is more risk averse than the former. Such a disparity in risk preferences makes trade possible between the bank and the customer, involving the bank selling the borrower a loan commitment that reduces uncertainty regarding the customer's future borrowing cost. With a variable-rate loan commitment, the bank still bears some interest rate risk but less than with a fixed-rate commitment. In essence, with a fixed-rate commitment the bank bears both the risk of changes in the index rate as well as of changes in the borrower's credit risk

7. This observation is due to Greenbaum, Kanatas, and Venezia (1991).

8. Loan commitment demand based on optimal risk-sharing considerations was first formally proposed by Campbell (1978) and later examined by Thakor and Udell (1987) to rationalize commitment and usage fees in loan commitment contracts. Also see Hawkins (1982), Holmstrom and Tirole (1993), James (1982), and Melnik and Plaut (1986).

premium, whereas with a variable-rate commitment the bank bears only the latter risk. In either case, the risk-averse borrower is transferring (some) interest rate risk to the bank, and to the extent that the bank is willing to participate at a price that is acceptable to the borrower, we have an explanation for why loan commitments are demanded by the bank's customers.

(b) Moral Hazard: One drawback of the previous explanation is that many loan commitment customers are large, publicly owned firms with numerous shareholders. From portfolio theory we know that even risk-averse shareholders should be indifferent to firm-specific (idiosyncratic) risk because they can diversify it away. Moreover, it is not clear why shareholders of nonbank firms should collectively demand a higher premium for bearing systematic risk than the bank's shareholders do.[9] So we would like to know if there will be a demand for loan commitments even when the bank's customers are not motivated by the desire to purchase insurance against interest rate risk.

One possibility is that loan commitments are effective in deterring *moral hazard*. The source of the moral hazard may be an inventive on the borrower's part to undersupply productive effort (relative to the case in which the borrower self-finances) or switch projects (in an undetected manner) to the bank's detriment. The intuition is as follows. We know from our discussions in Chapter 6 that the loan interest rate is distortionary in the sense that the higher this rate, the lower is the net return accruing to the borrower, and hence the greater is the borrower's incentive to reduce effort and/or switch to a riskier project. The consequences can be costly—the borrower may either need to post collateral or in extreme circumstances the bank may ration credit. A loan commitment provides a means for the bank to circumvent the distortionary effect of the loan interest rate without relying on more costly alternatives. This can be achieved by lowering the interest rate on the loan to a level sufficient to eliminate (or significantly diminish) moral hazard. This will generally mean that the bank will suffer an expected loss on the loan made under the commitment. This loss can be recouped through the commitment fee paid by the borrower at the time the commitment is made. The key is that the customer views the commitment fee as a sunk cost after it is paid, and hence the commitment fee does not affect either level of effort or choice of project. In this way the loan commitment helps to overcome moral hazard. The following example illustrates the point.

Example 8.1 Supose the management of Knight Apparel Company knows at $t = 0$ that it will have available at $t = 1$ an opportunity to invest \$100 in a risky project that will pay off at $t = 2$. Knight Apparel knows that it will be able to invest in one of two mutually exclusive projects, S or R, each requiring a \$100 investment. If Knight Apparel invests in S at $t = 1$, the project will pay off \$150 with probability 0.9 and zero with probability 0.1 at $t = 2$. If Knight Apparel invests in R at $t = 1$, the project will pay off \$158 with probability 0.7 and zero with probability 0.3 at $t = 2$. Knight Apparel's project choice is not observable to the bank from which it seeks to borrow the \$100.

9. Ignore for the time being the *risk-seeking* incentives provided to the bank's shareholders by the bank's access to a lender-of-last-resort facility and deposit insurance. We wish to focus for now on the possible economic motives for loan commitments, *abstracting* from the facilitating influence of regulation.

The riskless, single-period interest rate at t = 0 is 10 percent. It is not known at t = 0 what the riskless, single-period interest rate at t = 1 will be, but it is common knowledge that this rate will be 5 percent with probability 0.5 or 15 percent with probability 0.5. Assume universal risk neutrality and that Knight Apparel has no assets other than the project on which you (as the lender) can have any claim. Figure 8.3 depicts these data.

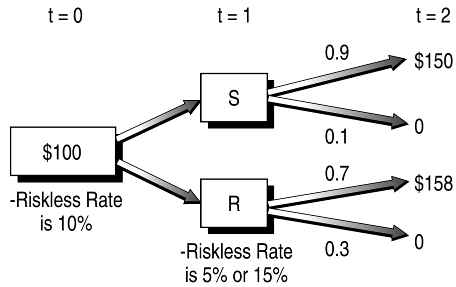


FIGURE 8.3 Investment Opportunities for Knight Apparel

Suppose you are Knight Apparel's banker and you know that Knight has two choices: (i) It can wait until t = 1 and then borrow in the spot market or (ii) it can purchase a loan commitment that will permit it to borrow at predetermined rates at t = 1. What advice would you give Knight Apparel? Assume a competitive loan market in which banks earn zero expected profits.

Solution We solve this problem in six steps. First, we consider alternative (i) and show that it is a Nash equilibrium for Knight Apparel to choose S at t = 1 if the spot riskless rate then is 5 percent. Second, we continue with alternative (i) and show that this Nash equilibrium fails to exist if the spot riskless rate at t = 1 is 15 percent. The reason is that the high interest rate diverts "too much" of Knight Apparel's cash flow into repaying the bank loan, so that the borrower prefers to gamble on the riskier investment R which, despite its lower success probability, gives Knight Apparel a higher *net* payoff in the successful state. The bank must therefore price the loan under the assumption that R will be chosen. But then the interest rate is so high that Knight Apparel declines the loan. Third, we point out that passing up the investment opportunity in the high-interest-rate state is socially wasteful because S has a positive total NPV even when the riskless rate is 15 percent. Fourth, we consider alternative (ii), and design a loan commitment contract that induces Knight Apparel to invest in S regardless of the spot riskless rate. Fifth, we solve for the commitment fee so that the bank can earn (at least) zero expected profit on the loan and the loan commitment taken together. Finally, in step 6 we calculate the *net* benefit of the loan commitment to Knight Apparel and show that it is positive.

Step 1 Consider alternative (i). Suppose the interest rate at $t = 1$ is 5 percent and you assume that Knight Apparel will choose S. Then the interest rate, i_s, that you should charge the borrower in order to just break even on the loan is obtained as a solution to the following equation:

(*Continued*)

$$0.9 \times (1 + i_s) = 1.05 \qquad [8.1]$$

where 0.9 is the probability that you will be repaid by Knight Apparel. Solving (8.1) gives $i_s = 16.67$ percent. If Knight Apparel chooses *S*, its expected payoff at $t = 2$ is

$$0.9 \times (150 - 116.67) = \$30 \text{ approximately.}$$

On the other hand, if Knight Apparel chooses R, its expected payoff at t = 2 is

$$0.7 \times (158 - 116.67) = \$28.93.$$

Thus, Knight Apparel will prefer S to R, and it is a Nash equilibrium for you to offer a \$100 loan at 16.67 percent.[1]

Step 2 Now suppose the interest rate at t = 1 is 15 percent. If you assume that Knight Apparel will choose S, then you should charge an interest rate, i'_S, that solves the following equation

$$0.9 \times \left(1 + i'_S\right) = 1.15. \qquad [8.2]$$

Solving (8.2) yields $i'_S = 27.78$ percent. If Knight Apparel does indeed choose *S*, its expected payoff at t = 2 will be

$$0.9 \times (150 - 127.78) = \$20 \text{ approximately.}$$

On the other hand, if Knight Apparel chooses R, its expected payoff at t = 2 will be

$$0.7 \times (158 - 127.78) = \$21.15.$$

Clearly, Knight Apparel will prefer R to S, and it is *not* a Nash equilibrium for you to offer the loan at 27.78 percent. But suppose you assume that Knight Apparel will choose R. Then, the interest rate, i'_R, that you should charge solves

$$0.7\left(1 + i'_R\right) = 1.15,$$

which yields $i'_R = 64.29$ percent. However, at this interest rate Knight Apparel will not borrow since its repayment obligation would exceed the maximum cash flow of the project.

Step 3 What this implies is that if Knight Apparel can only borrow in the spot market, it will invest only if the risklesss rate at t = 1 is 5 percent. If the rate is 15 percent, Knight Apparel will pass up its investment opportunity. This is a distortion in the following sense. Even when the riskless interest rate is 15 percent, project S has a positive total NPV, even though its NPV to Knight Apparel's shareholders is not positive. If Knight Apparel could somehow convince a bank that it would choose S if given a loan, the bank would be willing to extend the loan at terms that would enable the bank to break even and leave Knight Apparel with a positive NPV. However, credible communication from Knight Apparel to the bank may not always be possible (we have assumed that it is not), and if it is not, the bank must anticipate that Knight Apparel will act in its own best interest. The consequence is a bank loan that Knight is unwilling to accept, and a social waste represented by the foregone positive NPV of project S.

Step 4 We will now show that a loan commitment, negotiated at $t = 0$, can avoid this moral-hazard-induced loss. Suppose that under arrangement (ii), you offer to lend Knight Apparel \$100 (if Knight Apparel wishes to take the loan) at $t = 1$ at an interest rate of 16.67 percent, regardless of the spot riskless rate at that time. This is a fixed-rate loan commitment. As our analysis thus far has indicated, Knight Apparel will opt for S under these terms, so that your bank will break even on the loan if the riskless rate at $t = 1$ is 5 percent. Of course, if the riskless rate is 15 percent, you will lose money on the risky loan since you should be charging an interest rate of 27.78 percent in that case.[2] To recoup this loss, you should charge Knight Apparel a commitment fee at t = 0. What should this commitment fee be?

Step 5 To answer this question, note that your bank's loss, in terms of the amount that should be repaid in the successful state minus the amount that is actually repaid in the successful state, is

$$\$127.78 - \$116.67 = \$11.11.$$

The bank suffers this loss at $t = 2$ only if Knight Apparel's project succeeds (the bank also suffers a loss if Knight Apparel's project fails, but in that state the bank recovers nothing in either case), and the probability of success is 0.9. Hence, the bank's expected loss is

$$0.9 \times \$11.11 = \$9.999.$$

Since the probability of the 15 percent interest rate is 0.5 and we must discount from $t = 2$ back to $t = 1$ (at 15 percent) and from $t = 1$ back to $t = 0$ (at the 10 percent riskless rate prevailing at $t = 0$), we have the following present value at $t = 0$ of the bank's expected loss at $t = 2$

$$\frac{0.5 \times \$9.999}{1.15 \times 1.10} = \$3.95.$$

Thus, the commitment fee that the bank should charge Knight Apparel is \$3.95, given a zero expected profit on the loan *and* the loan commitment. It is important to note that Knight Apparel pays the commitment fee at $t = 0$, so that when it confronts its project choice at $t = 1$ it treats this fee as a sunk cost and its project choice is not affected by it.

Step 6 We can compute the overall benefit from the loan commitment by comparing Knight Apparel's NPV under arrangements (i) and (ii). Under (i), since borrowing

(*Continued*)

only takes place when the spot riskless rate at $t = 1$ is 5 percent, Knight Apparel's NPV is

$$\frac{0.5 \times \$30.00}{1.05 \times 1.10} = \$12.99,$$

where you will recall that \$30 is Knight Apparel's net expected payoff at $t = 2$ when it chooses S and is obliged to repay the bank \$116.67 (an interest rate of 16.67 percent). Under (ii), the expected NPV is

$$\underset{\substack{\text{(Riskless rate at}\\ t=1 \text{ is 5 percent)}}}{\frac{0.5 \times \$30.00}{1.05 \times 1.10}} + \underset{\substack{\text{(riskless rate at}\\ t=1 \text{ is 15 percent)}}}{\frac{0.5 \times \$30.00}{1.05 \times 1.10}} - \underset{\text{(commitment fee)}}{\$3.95}$$

$$= \$20.90.$$

Thus, Knight Apparel experiences a *net* gain of \$7.91 (which is \$20.90 – \$12.99) by purchasing the loan commitment. Note that this improvement is net of the commitment fee.

1. Note that this Nash equilibrium is not unique. If you assume that the borrower will choose R, then the loan interest rate you should charge is the solution to $0.7(1 + i_s) = 1.05$, which yields $i_R = 50$ percent. Now the borrower's expected payoff at $t = 2$ form choosing S is zero and its expected payoff at $t = 2$ from choosing R is $0.7(158 - 150) = \$5.6$. Thus, the borrower strictly prefers R to S, and it is also a Nash equilibrium for you to offer the loan at 50 percent. However, the Nash equilibrium we have focused on (that is, one involving a 16.67 percent interest rate) is strictly preferred by the borrower (lower interest rate and strictly higher expected rate) and you, as the lender, are indifferent because you make zero expected profit in each Nash equilibrium. Thus, competition among banks will ensure that the Nash equilibrium involving the 16.67 percent loan interest rate will prevail.

2. Note that the 27.78 percent interest rate is the correct breakeven rate for banks when the borrower chooses S. This assumption is validated now since the borrower will assuredly choose S when faced with a borrowing rate of 16.67 percent.

In this example, the loan commitment was useful in overcoming the moral hazard created by the possibility of undetected asset substitution by the borrower. A similar argument works for "effort aversion" moral hazard, and it suggests that loan commitments add value for borrowers; this observation has empirical support in that firms that purchase bank loan commitments experience abnormally positive stock price reactions upon announcing these purchases.[10] The conclusion is that borrowers may demand loan commitments because they are able to borrow on better terms under commitments than they could in the spot market. Banks are able to provide

10. See Shockley and Thakor (1997). The ability of the loan commitment to overcome moral hazard problems has been noted in Boot, Thakor, and Udell (1987, 1991) and Boot and Thakor (1991). A somewhat different rationale for loan commitments appears in Maksimovic (1990). Kanatas (1987) shows that loan commitments can convey information. Thakor (1989) highlights the role of commitments in resolving informational problems. Shockley (1992) shows that loan commitments reduce the agency costs of nonbank debt and hence facilitate an increase in the borrower's *total* leverage. See also James (1981) and Greenbaum, Kanatas, and Venezia (1991).

better terms because loan commitments avoid some of the moral hazard problems that plague spot loans.

(c) Liquidity Guarantee for Other Creditors: When a firm purchases a loan commitment, suppliers of inputs to the firm know that the firm will have access to liquidity equal to the amount of the commitment. This may reassure suppliers that the firm will have the funds necessary to service its debt obligations to them. Consequently, these suppliers may be willing to provide inputs to the firm on better terms than in the absence of the loan commitment. The result would be an overall lowering of the firm's cost of debt, which would benefit the firm's *shareholders*. This intuition can be seen in the following example.

Example 8.2 Suppose Northwestern Business Machines (NBM) has the opportunity to invest \$100 at $t = 1$ in a project that will yield a random payoff at $t = 2$. At $t = 0$ the firm is uncertain about the probability distribution of the random payoff of the project; this distribution depends on a state of nature, call it θ, that will be revealed privately (that is, it is not known to the creditors) to the management of NBM at $t = 1$ prior to making its decision of whether to invest in the project. At $t = 0$, all that NBM knows is that there is a 0.5 probability that $\theta = \text{G}$ (the "good" state), in which case the project will pay off \$200 with probability 0.9 and zero with probability 0.1. If the bad state occurs, the project will pay off \$130 with probability 0.9 and zero with probability 0.1.

At $t = 0$, NBM needs to buy \$20 of raw materials and other inputs if it is to proceed with the project at $t = 1$. The suppliers have agreed to provide trade credit so that the \$20 plus the agreed upon interest can be paid at $t = 2$.

The riskless interest rate that will prevail from $t = 1$ to $t = 2$ is 5 percent, and this is known to all at $t = 0$. Assume that the time that will elapse from $t = 0$ to $t = 1$ is so short that discounting can be ignored. Also assume that NBM's chief executive officer (CEO) will sustain a nonpecuniary cost (say the cost of personal effort) in initiating the project, and the pecuniary present-value equivalent of this cost is 1 dollar. All of the data for this problem are shown in Figure 8.4.

Compute the terms of trade credit as well as the NPV to NBM (net of the CEO's personal cost) if it: (i) borrows the \$100 in the spot credit market after learning θ, and (ii) purchases a loan commitment at t = 0 (prior to knowing θ) that would entitle it to borrow the \$100 at $t = 1$.

Assume that the bank as well as trade creditors provide credit at competitive terms, and that everybody is risk neutral.

Solution We solve this problem in five steps. First, we consider the spot credit alternative. We show that if suppliers price their trade credit to NBM at $t = 0$ assuming that NBM will undertake the project at $t = 1$ regardless of θ, then the project is undertaken only if $\theta = \text{G}$. Second, we argue that this suggests that the only Nash equilibrium is for trade creditors to believe that NBM will undertake the project at $t = 1$ if $\theta = \text{G}$ and not otherwise. We verify in the second step that this is indeed a Nash equilibrium. Third, we consider the loan commitment alternative. We show that there exists a fixed-rate loan commitment that induces NBM to invest at $t = 1$

(Continued)

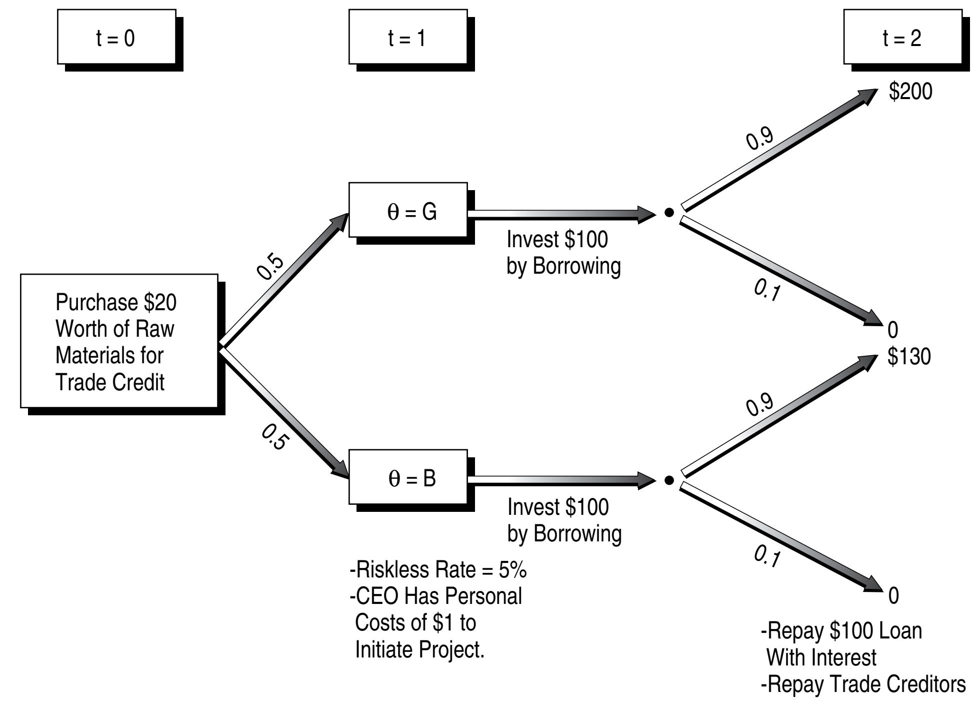


FIGURE 8.4 Investment Opportunities for Northwestern Business Machines

regardless of θ. Fourth, we verify that this commitment is part of a Nash equilibrium by checking that NBM will indeed prefer to purchase a loan commitment at $t = 0$ as opposed to borrowing in the spot market at $t = 1$. Finally, in step 5 we show that the loan commitment makes NBM better off ex ante because it lowers NBM's overall cost of credit. It achieves this by eliminating an underinvestment problem that results in trade credit being available to NBM at a lower cost than with spot borrowing.

Step 1 Let us first consider the spot credit alternative. Since the project has a success probability of 0.9 (regardless of θ) and the riskless rate is 5 percent, we know from Example 8.1 that the competitive loan interest rate is 16.67 percent. Thus, the repayment obligation of the \$100 spot loan will be \$116.67 percent. If the trade creditors assume that NBM will invest in the project regardless of θ, then the interest rate on trade credit should also be 16.67 percent, that is, the firm's repayment obligation should be $\$20 \times 1.1667 = \$23.33$. The total repayment obligation is then $\$116.67 + \$23.33 = \$140$. But then if $\theta = \mathrm{B}$, its NPV is

$$\frac{0.9(0)}{1.05} - \$1 = -\$1$$

where we have subtracted the decision maker's personal cost of \$1 in computing the NPV. Note that the zero in the numerator of the first term in the left-hand side of the above equation reflects limited liability (without which the zero would be replaced by

$130 - 140 = -\$10$). Thus, the firm will *not* undertake the project if it observes $\theta = \text{B}$. If $\theta = \text{G}$ is observed at $t = 1$, then the firm's NPV is

$$\frac{0.9(\$200 - \$140)}{1.05} - 1 = \$50.43$$

so that the project will be undertaken.

Step 2 This means that it cannot be a Nash equilibrium for the trade creditors or the bank to believe that NBM will undertake the project regardless of θ. Suppose they assume that the project will be undertaken only if $\theta = \text{G}$. Then, since the probability of the project being undertaken is only 0.5, the interest rate $\mathrm{i_t}$, charged for trade credit (with spot borrowing of the \$100 investment) must satisfy

$$0.5 \times 0.9 \times (1 + i_t) = 1.05,$$

which yields $i_t = 133.33$ percent. NBM's repayment obligation to the trade creditors is therefore $\$20 \times 2.3333 = \$46.67$. Its total repayment obligation now becomes $\$116.67 + \$46.67 = \$163.34$. If $\theta = \text{G}$ occurs, NBM's NPV is

$$\frac{0.9(\$200 - \$163.34)}{1.05} - 1 = \$30.42,$$

so that the project will be undertaken in that state. However, the project will not be undertaken if $\theta = \text{B}$. Hence, the beliefs of the trade creditors about NBM's future behavior are consistent (rationalized by NBM's behavior), and it is a Nash equilibrium for them to offer trade credit at an interest rate of 133.33 percent. Note that the NPV to NBM from the spot borrowing alternative, computed at $t = 0$ (prior to the realization of θ) is $0.5 \times 30.42 = \$15.21$, since the firm knows that it will invest only if $\theta = \text{G}$.

Step 3 Now consider the loan commitment alternative. Suppose NBM can obtain a loan commitment at $t = 0$ to borrow \$100 at 5 percent (the current riskless rate) at $t = 1$. Thus, its repayment obligation to the bank, if it borrows under the commitment, will be \$105. Since the repayment obligation that permits the bank to just break even is \$116.67, the bank's loss is $\$116.67 - \$105 = \$11.67$, and so the commitment fee (using the logic employed in Example 8.1) should be $\dfrac{0.9 \times 11.67}{1.05} = \$10.00$. Will the firm now invest in the project if $\theta = \text{B}$ at $t = 1$? To answer this, calculate NBM's NPV as $\dfrac{0.9(\$130 - \$105 - \$23.33)}{1.05} - 1 = \$0.43$, so that the firm will undertake the investment. Note that we have assumed here that trade creditors believe that NBM will undertake the project regardless of θ, if it has purchased this loan commitment at $t = 0$. As our analysis indicates, this assumption is warranted.

Step 4 To verify that we have a Nash equilibrium with the loan commitment, we also need to check that NBM will prefer to purchase the loan commitment at $t = 0$ as opposed to borrowing in the spot credit market.

The NPV for NBM, assessed at $\mathrm{t} = 0$, is

$$\frac{0.5[0.9 \times (\$200 - \$105 - \$23.33)]}{1.05} + \frac{0.5[0.9 \times (\$130 - \$105 - \$23.33)]}{1.05} - 10 - 1 = \$20.43.$$

(*Continued*)

Since this exceeds NBM's NPV with spot borrowing ($15.21), we have a Nash equilibrium with NBM purchasing a loan commitment at $t = 0$ to borrow $100 at 5 percent, and suppliers extending trade credit at 16.67 percent.

In this case the loan commitment reduces the firm's overall cost of credit. Recall from Chapter 5 that collateral solved an underinvestment problem. Here the loan commitment serves a similar purpose.[1]

Step 5 The underinvestment problem arises because the project has a positive NPV to NBM (and a positive total NPV) in state $\theta = \mathrm{B}$ only if the project's cost is just the $100 investment. In this case the *total* NPV of the project is

$$\frac{0.9 \times \$130}{1.05} - \$100 - \$1 = \$10.43,$$

so that the loan commitment helps to avoid a real underinvestment problem. This conclusion is appropriate if we view the $20 worth of raw materials as a purchase that NBM would make even if the project were not available, that is, the raw materials do not add to the cost of the project. But if we interpret that $20 as adding to the cost of the project (that is, these raw materials would not be purchased if the project were unavailable), then the total NPV of the project is

$$\frac{0.9 \times \$130}{1.05} - \$100 - \$20 - \$1 = -\$9.57.$$

In this case the project is *socially inefficient* in the $\theta = \mathrm{B}$ state, so that the loan commitment (which still results in the project being undertaken when $\theta = \mathrm{B}$) does *not* resolve an underinvestment problem in the usual sense.[2] Indeed, it ends up inducing an *overinvestment*[3] by NBM that makes it better off *ex ante*. In this case the role played by the loan commitment[4] is quite different from that played by collateral in our discussions in Chapter 5.

1. Berkovitch and Greenbaum (1990) have suggested that a loan commitment can eliminate underinvestment. The example presented above captures their intuition.
2. To reiterate, by an "underinvestment problem" we mean a situation in which the firm passes up a project with positive total NPV.
3. That is, the firm invests in the project when its total NPV is negative. It does so because the NPV to its own shareholders is positive.
4. In both this illustration as well as in Example 8.1, we have assumed that the firm has the liquidity to pay the commitment fee on its loan commitment. Why would the firm not want to use this liquidity to provide equity to the project instead and thereby reduce its total borrowing? Boot, Thakor, and Udell (1987) examine this issue and show that the borrower is *strictly better off* using its liquidity to pay the commitment fee rather than using it as equity in conjunction with a spot loan.

Borrowers often use loan commitments as an assurance to other creditors. For example, commercial paper borrowers routinely purchase dedicated bank loan commitments explicitly to back up commercial paper issues.

(d) Protection Against Future Credit Rationing: A borrower's future access to credit is threatened by three possibilities: (1) deterioration in its own credit rating, (2) deterioration in the general market availability of credit, and (3) changes in bank-specific factors that diminish the bank's ability to provide credit. A loan commitment may protect the buyer against the first two possibilities.[11] Of course, the MAC clause in the loan commitment contract limits the usefulness of the commitment as insurance against rationing due to the first possibility. However, the empirical evidence suggests that this motive is present in the purchase of many loan commitments.[12] Interestingly, theory suggests that the interaction between this motive for purchasing loan commitments and the MAC clause can cause banks to *overlend* under commitments when the economy is doing well.[13]

A Federal Reserve survey (1988) sheds light on the reasons for loan commitment demand.[14] The most frequently mentioned reasons given for commitments were "general convenience and minimizing loan arrangement costs," and "protection against general credit crunches." The next most frequently mentioned reasons were to "ensure credit access against a creditworthiness deterioration" and "to lock in a fixed markup over a reference interest rate."

(e) Reducing Market Incompleteness: When the capital market is *incomplete* (recall our discussion in Chapter 1), investors and firms lack all of the risk-sharing opportunities they desire. Thus, if the market is incomplete and the loan commitment produces a payoff stream for the borrower that cannot be replicated by linear combinations of existing securities (as, for example, in Example 8.1), then the availability of a loan commitment reduces market incompleteness. Since investors now have access to expanded risk-sharing opportunities because they can invest in firms that purchase loan commitments (as well as those that do not), these investors may be made better off by the availability of loan commitments. In other words, there may be a demand from investors for payoff patterns that can only be produced by firms that purchase loan commitments.

Pricing of Loan Commitments

The Model

The Analogy Between Loan Commitments and Options: We develop an approach for pricing loan commitments, based on the observation that their payoff structure resembles that of a common stock put option.[15] As discussed in Chapter 1, a put option is the right to sell a security (the deliverable) at a fixed price during some fixed time interval, or at some fixed future date. The major components of this contract are the:

i) identity of the deliverable,
ii) option price,
iii) strike price, and
iv) exercise date or period.

11. Morgan (1989) has developed a model to show how a commitment can solve a credit rationing problem. See also Glick and Plaut (1989).

12. See Sofianos, Wachtel, and Melnik (1990), and Berger and Udell (1990). Thakor (2005) develops a theoretical model that explains how loan commitments can protect borrowers against credit rationing despite the presence of the MAC clause.

13. See Thakor (2005).

14. See Avery and Berger (1991).

15. This observation is due to Thakor, Hong, and Greenbaum (1981), and Thakor (1982).

For example, \$500 might be paid for the right to put (sell) 100 shares of General Motors common stock at \$50/share (or \$5,000) at any time over the next six months. The option price is \$500, the strike price is \$50/share, the deliverable is 100 shares of General Motors common stock, and the exercise dates are all dates extending over the next six months. The "writer" of the option accepts the \$500 option price in exchange for the responsibility to purchase 100 shares of GM stock for \$5,000 at the discretion of the option buyer at any time during the next six months. (Some options are exercisable only at the end of the term, rather than at any time during the term.)

Now consider the bank loan commitment. The loan commitment buyer pays a commitment fee (option price) for the right to put (sell) a security to the bank at a prespecified price over some pre-established time interval. The security is the commitment owner's IOU (debt) and the strike price is the face (par) value of the loan, that is, the dollar amount of the borrowing. The time interval is the life of the commitment. Hence, in selling loan commitments, banks are writing put options where the underlying deliverable is the debt instrument of the commitment buyer. The commitment buyer will take down the commitment (exercise the put option) if the value of its debt instrument on the exercise date is less than the committed loan amount (the strike price). The difference between the loan amount and the debt instrument value at the time of commitment exercise represents the customer's gain from exercising the commitment, and the present value of this gain at the time of commitment purchase should be the commitment fee or price the customer is willing to pay.

The Model: Suppose we wish to value a loan commitment issued at $t = 0$ that would allow the purchaser to borrow \$F (the face value of the loan or strike price of the put option) at $t = 1$ at some predetermined interest rate i_c. The maturity of the loan will be one period, that is, it will mature at $t = 2$, and the loan (if taken) will be free of default risk. Assume that the current one-period riskfree rate is i_o. Assume that the one-period yield on the borrower's debt at $t = 1$ will either be $i_1^+ > i_o$ or $i_1^- < i_o$.

The probability of i_1^+ is p and the probability of i_1^- is 1 − p. Assume $i_1^+ > i_c > i_1^-$. Everybody is risk neutral. What is the value of this fixed-rate commitment?

Solution: At $t = 1$, suppose the spot yield on the borrower's debt is i_1^-. Then it is clear that the borrower has no incentive to take down the loan commitment since cheaper credit is available in the spot market. But if the spot yield is i_1^+, then the borrower will take down the commitment since the commitment rate is $i_c < i_1^+$. The value of the borrower's debt at $t = 1$ in this state is:

$$\frac{F[1 + i_c]}{[1 + i_1^+]} \tag{8.3}$$

where $F[1 + i_c]$ is the borrower's future repayment obligation at $t = 2$, which is discounted back to $t = 1$ at the spot yield i_1^+. Note that the borrower is receiving \$F from the bank when it takes down the loan, and in exchange the bank is receiving a debt security worth the amount given by [8.3]. That is, the borrower is selling the bank a debt security worth $F[1 + i_c]/[1 + i_1^+]$ for \$F when it exercises its loan commitment put option. The gain to the borrower from exercising the put option is:

$$F - \frac{F[1 + i_c]}{[1 + i_1^+]} \tag{8.4}$$

The value of the loan commitment to the borrower at $t = 0$ is then:

$$\frac{p\left\{F - \frac{F[1+i_c]}{[1+i_1^+]}\right\}}{[1 + i_0]}$$

where the expression in [8.4] is multiplied with the probability of the yield i_i^+ and is discounted back to $t = 0$ at the riskless rate i_o (since everybody is risk neutral).

We have thus far discussed the valuation of fixed-rate loan commitments. Variable-rate commitments can be valued similarly. The add-on to the index rate in the variable-rate commitment would be held fixed. However, this add-on is a premium charged by the bank for the customer's default risk in excess of that reflected in the index rate. Thus, the customer will exercise the commitment whenever the fixed add-on is smaller than the add-on the customer would be charged in the spot market. Once again, we have a put option purchased by the commitment buyer. The difference is that with a fixed-rate commitment the customer is purchasing protection against an increase in its total borrowing cost (which includes an increase in the index rate as well as in the add-on reflecting borrower-specific risk), whereas with a variable-rate commitment the customer is purchasing protection only against an increase in the add-on due to a decline in its own credit rating.

Empirical Predictions of Valuation Model

The valuation model developed above suggests that borrowers purchase loan commitments to lock in borrowing rates. Hence, more commitments should be exercised when borrowers experience an increase in their cost of spot-market borrowing.

There is abundant anecdotal evidence to support this prediction. For example, in 1990, Travelers Corporation, a Hartford-based insurance company, drew down a substantial portion of its \$1.075 billion credit line after the major rating agencies downgraded its credit rating (and thereby increased its cost of borrowing in the spot credit market). It was reported that the company sought to ensure liquidity and assure its access to short-term funding after boosting loan-loss reserves by \$650 million.[16]

Another testable prediction of the valuation model is that the cost of loan commitments should increase as the volatility (future uncertainty) of the customer's spot borrowing rate increases. This prediction follows immediately from the well-known property of put options that they increase in value as uncertainty in the future value of the underlying asset increases.

The Differences Between Loan Commitments and Put Options

While there is a striking similarity between a common stock put option and a bank loan commitment, there are also important differences. Four key differences are as follows:

16. Lipin (1990) reported that, "The move by Travelers is not expected to be an isolated event. More corporations will seek to maintain liquidity in a precarious economic environment, lenders say. In addition, the bank lines have become more attractive due to rising rates and other problems in the commercial paper market."

(1) An exchange-traded put option is a binding contract—the option seller is legally liable for the contractual payment if the option is exercised. By contrast, due to the MAC clause, a bank loan commitment is a discretionary contract.

(2) An exchange-traded option is a *transferable* contract, whereas a loan commitment is not. That is, if firm A buys a loan commitment from a bank, it cannot sell this commitment to firm B—loan commitments are not transferable. The commitment owner may of course exercise the commitment and lend the proceeds to firm B, but this is yet a different transaction.

(3) Loan commitment pricing differs from that of exchange-traded options. For example, a loan commitment may include a usage fee that is an increasing function of the unused portion of the line. This is inconsistent with the option pricing formulation. One way to understand usage fees is in the context of the earlier explanation that banks may offer loan commitments because they provide information about future loan demand. Deviations of actual takedowns from expected takedowns under commitments represent prediction errors that may be costly to the lender. For example, if the lender incurs a cost of preparation (funding) to make the expected loan and therefore finds it costly to invest the planned funds in something other than the loan, then the lender's cost increases with the *error* in takedown prediction. Assessing a fee on the unused portion of the commitment is a way to induce the customer to provide the lender more accurate information about future loan demand.[17]

(4) A put option is either exercised in full, or not at all. Loan commitments typically do not exhibit such takedown behavior. Loan takedowns, F^*, are usually only some fraction of F, the face value of the commitment. There are two possible explanations for this partial takedown phenomenon:

   (i) The customer lacks the "need" for all of the funds that can be borrowed under the commitment.

   (ii) The customer has a long-term relationship with the bank and seeks to foster good relations by not fully exploiting windfalls.

Consider (i) first. Its reasonableness depends on the customer's access to nonnegative NPV investment opportunities. If the customer's financial leverage is unrestricted and it has unlimited investment opportunities that yield nonnegative NPVs, then we can expect its demand for funds to be highly elastic to its borrowing rate, and the commitment is likely to be exercised in full or not at all, as implied by the option valuation model. However, positive-NPV investment opportunities are typically limited.[18] Moreover, their ability/willingness to borrow under the commitment may be constrained by capital structure considerations, including restrictions imposed by covenants in outstanding debt contracts. In this case, loan demand will be imperfectly elastic to interest rates, and partial takedowns would then be possible. This is illustrated in Figure 8.5.

17. A related explanation appears in Thakor and Udell (1987) who propose that usage fees help the bank to sort out borrowers with high and low takedown probabilities when they are privately informed about their takedown probabilities and all request the same commitment amount.

18. Investing in marketable securities is typically a negative NPV alternative for corporations owing to transactions costs and taxes.

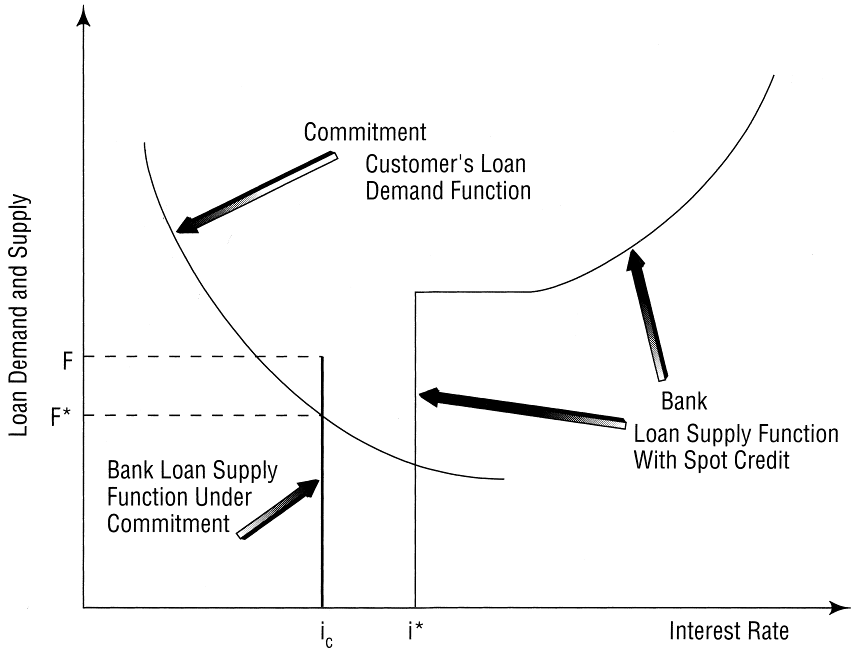


FIGURE 8.5 Partial Takedown With Imperfectly Elastic Loan Demand

When the customer's spot borrowing rate is i^* and its commitment rate is i_c, its loan demand is F^*, which is less than the credit line F. The bank's loan supply function under the commitment is a vertical line, indicating that the bank is willing to lend any amount up to F at i_c under the commitment. With a spot borrowing rate of i, the loan supply function *may* look like that indicated in the graph. This function says that the bank is willing to lend any amount up to some number (possibly) exceeding F, at a rate of i^*, and that the amount the bank is willing to lend may not increase for relatively small increases in the interest rate beyond i^*; for sufficiently higher rates, the bank may be willing to lend more.

Now consider (ii). A customer's takedown behavior may be seen as influencing the future pricing or availability of bank services.[19] This link presupposes some cost to the borrower of changing banks or incomplete exchange of information among banks. For example, information reusability will give the incumbent bank an advantage over competing banks with respect to information about the customer. This could enable the incumbent to offer credit at better terms than competitors could, thereby making it costly for the customer to switch to another bank. Now, since the customer's exercise of the commitment imposes a loss on the bank, it is reasonable to expect the bank to adjust its loan commitment pricing based on observed takedowns. For example, if the customer develops a reputation with the bank for taking down no more than 50 percent of its line of credit, the bank will begin to price the loan commitment taking that into account. This will yield a lower commitment price than if the customer took down 100 percent of the previous commitment. Alternatively, the bank will raise the commitment price if it expected the customer to take down 30 percent of the previous commitment and it actually took down 50 percent.

19. See Thakor, Hong and Greenbaum (1981), and Greenbaum and Venezia (1985).

TABLE 8.4 Similarities and Differences Between Loan Commitments and Put Options

	Put Option	Loan Commitment
Similarities	1) Deliverable or underlying security.	1) Customer's indebtedness (IOU).
	2) Option price.	2) Commitment fee.
	3) Strike price.	3) Size of the loan commitment (F).
	4) Exercise date.	4) Date commitment can be taken down.
Differences	1) Binding contract.	1) Discretionary contract.
	2) Transferable (tradeable) contract.	2) Nontransferable contract.
	3) No usage fee.	3) Usage fee.
	4) Exercised either in full or not at all.	4) Often partially exercised.

Of course, one could argue that the customer should *explicitly* reduce the size of the commitment if it does not plan to use all of it. However, the customer may still request a larger commitment than it needs under *normal* circumstances because of the possibility that an unexpectedly large credit need may arise in the future. But to the extent that the bank perceives that the probability of that happening is low, the price of the commitment will be lowered by the customer's previous partial takedowns. This phenomenon is similar to an automobile owner choosing *not* to file some auto collision claims with his insurance company due to the (adverse) learning the insurance company engages in when a claim is filed. In Table 8.4 we summarize the similarities and differences between loan commitments and put options.

Loan Commitments and Monetary Policy

Regulators conduct monetary policy by altering the quantity of credit or money supply and its *price* (interest rates). Loan commitments are a source of slippage in the Fed's ability to conduct monetary policy.[20] The reason is that once a commitment is sold, the amount of lending is determined by the customer's demand for funds at the prespecified interest rate. Now suppose the Fed wishes to implement a *contractionary* monetary policy. Using open market operations, the Fed would sell securities and drive up interest rates. While the higher interest rates reduce the demand for *spot* credit, they make borrowing under prearranged loan commitments more attractive and thereby increase takedowns.[21] Total bank lending may thus actually *expand* in the short-run in response to a contractionary monetary policy. This short-run perversity is likely to be reversed eventually as banks adjust by reducing the volume of their loan commitments in subsequent periods. Nevertheless, the growth of loan commitments can increase money market turbulence and frustrate monetary policy efforts.

20. This observation has been made by Deshmukh, Greenbaum, and Kanatas (1982), Duca and Van Hoose (1990), and Wojnilower (1980).

21. This is because an increase in market interest rates increases the cost of spot borrowing for the bank's customers, whereas the commitment rate either stays the same (under fixed-rate commitments) or rises less (under variable-rate commitments).

Other Contingent Claims: Letters of Credit

Loan commitments are not the only contingent claims that have registered striking growth in recent years. In this section we discuss two others that have grown impressively, commercial and standby letters of credit.

Commercial Letters of Credit and Bankers Acceptances

Commercial letters of credit (L/Cs) are used to facilitate trade, most commonly international, and are one of the oldest of banking contracts. In a typical transaction involving an L/C, the exporter has limited knowledge of the importer's ability to pay and limited ability to enforce contracts across national boundaries. The exporter therefore asks the importer to arrange for its bank to issue an L/C guaranteeing payment to the exporter upon presentation of the appropriate shipping documents. The exporter obtains the bill of lading and other shipping documents when goods are loaded on the ship for export. The L/C is a promise by the importer's bank to pay the exporter, given the necessary shipping documents. Thus, as the third party to the transaction, the bank substitutes its own creditworthiness for that of the importer and thereby reduces the default risk confronting the exporter.

When the exporter presents the necessary documents to the paying bank, it receives either a *sight draft* (immediate payment) or a *time draft* promising payment at some future date. In the latter case, the resulting instrument becomes a *bankers acceptance*, which is marketable and usually quite liquid. Thus, a bankers acceptance can be viewed as an outcome of a commercial L/C. Any draft "accepted" by a bank in the performance of its obligation under a commercial L/C is a bankers acceptance.[22]

In other words, a commercial letter of credit is essentially a *performance guarantee*. It can be defined as a promise to endorse or "accept" a time draft conditional on prespecified terms being satisfied. The act of accepting the time draft implies that, from the exporter's viewpoint, the bank's promise to repay replaces that of the debtor, and this creates a negotiable security. Consequently, the bank bears the risk that the debtor (importer) may default. Figure 8.6 depicts the steps leading to the creation of Banker's Acceptances. For simplicity, we have included only the importer's bank. Sometimes the exporter's bank is also involved as an intermediary between the exporter and the importer's bank, and time drafts may be accepted by both banks, giving rise to "two-name paper."

If the importer's (or buyer's) bank accepts a time draft and thereby creates a bankers acceptance, it has two choices. It can either hold the acceptance or it can sell it in the secondary market. If it decides to hold the acceptance, it ends up funding the credit (it has essentially extended a loan to the importer), so that the act of acceptance is automatic. However, if the acceptance is sold in the secondary market, the *holder* of the acceptance will provide funding, but the bank *guarantees* payment.[23]

22. See Melton and Mahr (1981).

23. When a bankers acceptance satisfies the purchase criteria laid down by the Federal Reserve Bank (Section 13 of the Federal Reserve Act), it is called "eligible" and can be sold to the Fed.

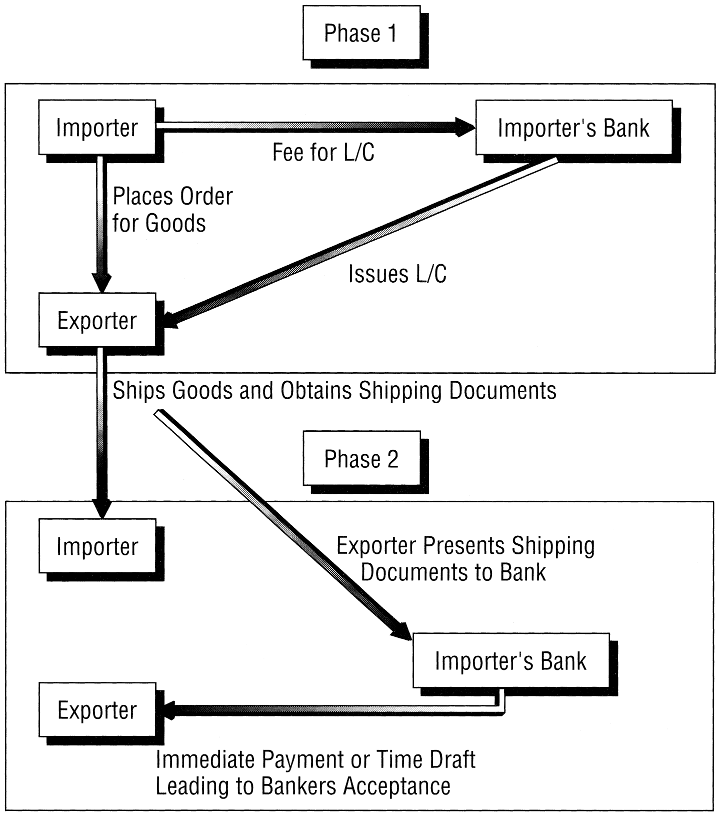


FIGURE 8.6 The Letter of Credit, Banker's Acceptance Nexus

Standby Letters of Credit

A *standby letter of credit* also guarantees the performance of an "account party," usually in a commercial or financial transaction, but it does not necessarily involve a funding transaction. A standby L/C issued by the "second party's" (the buyer or debtor or the party that owes some sort of performance to the "first party") bank obligates that bank to compensate the first party (the seller or the creditor or the party that is owed) in the event of a performance failure. The second party would then be liable to its bank for the disbursements the bank made under the L/C. From this perspective, standby and commercial L/Cs are similar. However, with a commercial L/C the issuing bank usually advances payment and is repaid by its customer, whereas with a standby L/C the bank makes payment only if its customer fails to fulfill a contractual obligation. Consequently, the bankers acceptances associated with commercial L/Cs have no counterpart among standby L/Cs.

Standby L/Cs are often used in international trade to facilitate transactions in which the seller has insufficient knowledge of the buyer's creditworthiness. Of course, the seller must still rely on the buyer's bank to "make good" on its promise, which is why there is often a second bank—typically the seller's—that augments the issuing bank's guarantee with its own. Such L/Cs are known as *confirmed letters of credit*.[24]

24. See Thakor (1988), and Greenbaum, Soss, and Thakor (1986).

Standby L/Cs are also used to guarantee performance in contracts involving greater variety and complexity than the simple international trade contract described above. Through standby L/Cs, banks now operate in areas that were once the exclusive domain of bonding, title, and insurance companies. For example, suppose a builder promises to deliver a completed building by a prespecified date or face a predetermined penalty. The buyer could ask the builder to obtain a standby L/C to guarantee the contract. Thus, if the builder fails to keep its promise, the buyer can collect the penalty amount from the bank that issued the L/C. The builder would then be responsible to pay its bank the penalty amount disbursed earlier by the bank. In banking, standby L/Cs are used as credit enhancements for securitizations and back-ups for commercial paper when the market gets skittish, that is, they replace loan commitments, thereby avoiding the risk of the MAC clause.

The Option-Like Feature of Standby Letters of Credit

Standby L/Cs can also be viewed as *put options*, like loan commitments. In the case of a loan commitment, the customer purchases an option to sell to the bank a security (the customer's indebtedness) that may be of less value at the time of exercise than the exercise or strike price (the amount loaned to the customer). In the case of a standby L/C, the bank agrees to purchase from the creditor a claim (the debtor's indebtedness) at par, contingent on the failure of the primary debtor to "perform," that is, to honor the claim. That is, the "second party" (the creditor) has the option to "put" the primary debtor's debt claim to the bank when nonperformance by the debtor renders the value of its debt claim less than par. In exchange for writing the option, the bank collects a fee. The option feature of a standby L/C implies that a bank that issues this instrument is conveying to the buyer a contingent claim and imposing on itself a contingent liability. The latter becomes an actual liability if the primary debtor fails to perform under the stipulations of a contract.

One important difference between loan commitments and standby L/Cs as put options is in the random processes influencing the market values of the underlying claims in the two cases and in the consequent trigger mechanisms giving rise to exercise. In the case of a loan commitment, an increase in the customer's spot borrowing rate, above the commitment. In the case of standby L/Cs, nonperformance by the debtor depresses the value of the claim below the strike price (the guaranteed value of the claim), prompting exercise of the option. Another important difference lies in enforceability. Unlike the loan commitment, the standby L/C does not have a MAC clause and is therefore more rigidly binding.

Other Contingent Claims: Swaps

What Are Swaps?

A *swap* is an agreement between two parties to exchange their exposure to a specific risk. The trade often involves an intermediary acting as either principal or broker.[25] Thus, for example, a swap is a tool for managing various types of risk.

25. See Antl (1983), Baecher (1991), Beidleman (1985), and Loeys (1985).

Basically an interest rate swap involves exchanging interest payments on notional securities with different prospects such as duration or the method by which interest payments are determined. For example, suppose a firm has a floating-rate liability and a fixed-rate asset. Such a firm will suffer losses if interest rates rise sharply. Now suppose another firm has a fixed-rate liability and a floating-rate asset. This firm will suffer losses if interest rates fall sharply. These two firms could arrange a swap to exchange their interest payments and thereby reduce their exposures to interest rate risk.

Interest rate swaps were first used in the Eurobond market during 1981. Large international banks, which lend mostly on a floating-rate basis, were the first to use swaps in which they exchanged the fixed-rate interest obligations on their liabilities for lower-cost floating-rate interest payments on equivalent notional amounts of claims. The swap market migrated to the United States in 1982 when the first domestic swap took place between Sallie Mae (Student Loan Marketing Association) and the ITT Financial Corporation. Since then this market has experienced explosive growth, and is now trillions of dollars of notional claims.

A typical swap involves \$25 to \$75 million of debt with 3- to 10-year maturity on one side of the transaction, and a floating-rate loan typically indexed to the LIBOR, the prime, or the T-bill rate on the other side.

How a Swap Works

Suppose we have two firms. Firm A is a bank with \$150 million of loans that promise a floating interest rate of prime plus 25 basis points, financed with \$150 million of 10-year bonds promising fixed 10 percent interest rate. Firm B is an S&L with \$150 million of fixed-rate mortgages financed with short-term MMFs (money market funds) and CDs with interest rates indexed to the T-bill rate. Each institution is exposed to interest rate risk that it wishes to hedge.

We could now arrange a \$150 million, 10-year interest rate swap between the bank and the S&L. The swap may be structured as follows. The S&L agrees to pay the bank a fixed rate of 10 percent per year on \$150 million, for 10 years. In return, the bank agrees to make the S&L a floating-rate payment at 2.5 basis points above prime, on a \$150 million principal. In this way the bank and the S&L have effectively exchanged their liabilities. Each has now hedged its interest rate exposure since the fixed-rate liability more closely matches the S&L's fixed-rate assets, whereas the floating-rate liability more closely matches the bank's floating-rate assets. Figure 8.7 depicts this arrangement.

Early on, swap transactions normally involved an intermediary functioning as a broker—typically a commercial bank or an investment banker.[26]

More recently, intermediaries have performed more like asset transformers, effectively providing guarantees to both parties to a swap transaction. For example, if the bank in the above transaction defaults, the intermediary would collect the fixed 10

26. For example, Loeys (1985) reports a \$100 million swap transaction, similar in nature to the one in our example above, in which the swap broker's fee was \$500,000.

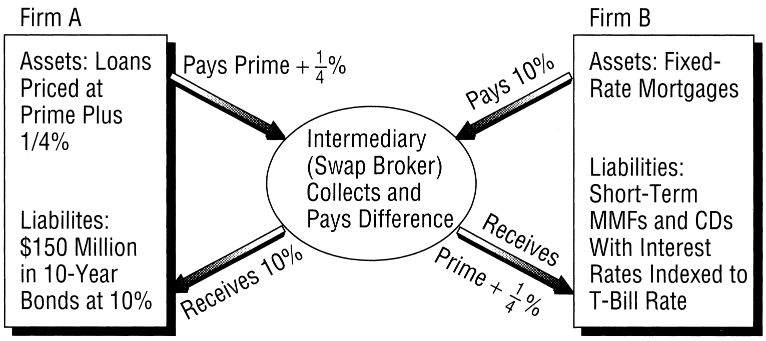


FIGURE 8.7 An Example of an Interest Rate Swap

percent from the S&L and make payments to it at 25 basis points above prime. Thus, *it would assume the role of the bank* until it can find an appropriate firm to replace the departed bank. And to the extent that it may not have the bank's balance sheet, the swap broker would expose itself to interest rate risk. For example, if interest rates were to rise sharply, the intermediary would lose.

Traditionally the most common type of swap was the one described in our example, namely that involving a dollar fixed-rate loan swapped for a dollar floating-rate loan. Such a swap is called a "plain vanilla" swap. Recently, however, different types of swaps have proliferated. One is a floating-to-floating swap where parties agree to swap floating rates based on different index rates. For example, a bank whose assets are floating-rate loans at prime plus 20 basis points and whose liabilities are floating-rate CDs at LIBOR minus 40 basis points may wish to swap the interest payments on its liabilities with those of an institution that has the interest rate on its liabilities indexed to the prime rate. Such swaps are known as *basis swaps*.

Another popular swap involves currencies. For example, a bank may have foreign loans financed by domestic deposits, so that the interest payments on its loans may be denominated in Japanese yen, while the interest payments on its deposits may be denominated in dollars. Such a bank might wish to swap its yen-denominated payments for dollar-denominated payments (perhaps with a Japanese bank that has dollar-denominated loans financed by yen-denominated deposits raised in Japan).

There are two common types of currency swaps: traditional *fixed/fixed currency swaps* and *cross-currency interest rate swaps*. A fixed/fixed currency swap involves fixed interest rates in each currency. Principal may or may not be exchanged. If principal is exchanged, this kind of swap transforms a fixed coupon bond denominated in one currency into a fixed coupon bond in another currency. With a cross-currency interest rate swap one exchange a fixed payments stream for a floating payment stream, as well as payments in different currencies. These contracts are occasionally combined in a single transaction, and sometimes the currency and interest rate components are separated. There are other variations as well. For example, there are swaps in which the two parties exchange yields on assets of different maturities (or currency denominations), rather than interest payments on liabilities. The point is that a swap can be tailor-made to suit the needs of the swapping parties, so that the potential variety of swaps is almost limitless. Some of these are discussed in the next subsection.

Swaps and Swap-Related Innovations

(a) Interest Rate Swap Variations: Some variations on the basic interest rate swaps are listed below.

- **Amortizing Swap:** This is a swap in which the notional principal amount diminishes over the life of the swap in a specified manner. This may be done so that payments match the expected cash flows of a financing project or the prepayment schedule of a mortgage.
- **Indexed Amortization Swap:** This is an amortizing swap in which the amortization of the notional principal depends on the stochastic value of some index like say the 3-month LIBOR.
- **Forward Swap:** This is a swap that does not begin until a designated future date. The fixed rate in the swap is linked to spot market rates, and the swap must be executed on the prespecified date.
- **Step Up/Down:** This is a swap in which the fixed-rate payments level varies, either increasing or decreasing over some portion of the swap term. For example, the fixed rate in the swap might be set below the market for the first 2 years, with an above-market rate for the remainder of the term.

(b) Swaps Involving Asset Payoffs Other Than Interest Rates

- **Commodity Swaps:** In a commodity swap, the contracting parties agree to exchange payments based on the value of a particular physical commodity, for example, gold, oil, or silver. One party pays a fixed price for the commodity and receives the spot price of the commodity at some future date. This relatively new contract that may appeal to commodity fund managers is generally short (2 to 3 years), but maturities up to 7 years are available.
- **Indexed Returns Swaps:** In this swap, one of the payments is linked to the total return of a market portfolio, like the S&P 500. This return can be exchanged for a payment stream based on either a fixed rate, such as the current T-bill rate (for a specific maturity) plus 30 basis points, or some floating rate (for example, the LIBOR). An interesting type of indexed return swap is a *foreign indexed swap*, which is designed to capture the relative performances of security types (for example, U.S. equities vs. Japanese equities). For example, suppose an investor owns a 5-year U.S. floating-rate note yielding LIBOR plus 50 basis points. This investor wants to invest in Japanese government bonds, but cannot trade the securities directly and wants to manage the foreign exchange risk. He can enter into a swap whereby he receives the dollar equivalent of the monthly returns on the Japanese bond and pays LIBOR, giving him a total return equal to the Japanese bond return plus 50 basis points.
- **Mortgage Swaps:** This is a swap that replicates all or a portion of the return characteristics of mortgage securities. In the most basic structure, a mortgage yield is exchanged for a floating-rate return, and the notional balance on which the payments are based is amortized according to either a specified schedule or the actual prepayment experience of the underlying pool of mortgages. The most recent innovation in this class is an *indexed amortization swap* in which a fixed-rate payment is exchanged for a floating-rate payment, but the notional balance amortizes according to a schedule that depends on the movements in

the yield of a prespecified security. For example, if the yield on the security falls by somewhere between 50 and 100 basis points, then the balance will amortize by 7 percent over the next period.

(c) Derivative Securities Based on Swaps

- **Swaptions:** With a swaption, one of the contracting parties has the option to allow an existing swap to be terminated or extended. These contracts are also called *cancelable, callable*, or *putable swaps*, and they can either be American or European in their options characteristics. Thus, a swaption is basically an option on a swap. Suppose two parties, A and B, enter into a contract in which A sells a *call swaption* to B. Then, at the exercise date, B can choose whether or not to exercise the option. If B exercise his option, he enters into a swap to receive, say, a fixed-rate payment in exchange for a floating-rate payment. These payment terms are all prespecified, as in a regular swap. The only difference is that one of the two parties has the legal right to decide whether or not to execute the swap at a future date.
- **Caps:** A *cap* is a swap contract in which the interest payments themselves have option characteristics. That is, the exercise (strike) price is set at particular interest rate levels. For example, suppose party A goes to a swap broker and buys a cap based on the 3-month LIBOR from a "cap writer" (party B), who represents the other party to the contract. Party A pays a premium (the price of the options) to the swap broker who subtracts his fee and passes along the remainder to party B. Now, party B is obliged to periodically (on each reset date) pay party A an amount equal to

$$\text{notional principal} \times \max\{0, \text{3-month spot LIBOR} - \text{strike rate}\},$$

where max (x,y) means the greater of x and y. Suppose the strike rate is 10 percent. Then if the 3-month spot LIBOR is 12 percent, party B must pay party A an amount equal to 2 percent of the notional principal, whereas if the 3-month spot LIBOR is 9 percent, party B pays nothing on the reset date. Thus, a cap is simply a sequence of consecutive expiration options. These options can be viewed as call options on the specified interest rate or put options on the underlying security. When rates rise, the security's price falls and the option becomes more valuable. As with a standard common stock option, the value of a cap (and hence the initial option premium) increases as the interest rate rises.

The cap market has developed numerous derivatives and customizations. Some of these are:

- **Floors:** Here party B pays party A an amount equal to notional principal $\times$ $\max\{0, \text{strike rate} - \text{spot market rate on a specific security}\}$ on the date of exercise (reset date) of the periodic option.
- **Collars:** Here party B pays A an amount equal to notional principal $\times$ $[\{0, (\text{spot rate} - \text{cap strike rate})\} - \max\{0,(\text{floor strike rate} - \text{spot rate})\}]$. That is, A is buying a cap from B and simultaneously selling a floor to B.

Suppose the cap strike rate is 15 percent and the floor strike rate is 10 percent .Then if the spot rate on the chosen security is 17 percent, the spot rate minus the cap strike rate is 2 percent, so party A receives 2 percent of the notional principal. If the spot rate is 9 percent, then the floor strike rate minus the spot rate is 1 percent, and party A receives 1 percent of the notional principal. If the spot rate falls between the cap and floor strike rates (say at 12 percent), then party A receives nothing.

Advantages and Disadvantages of a Swap as a Hedging Instrument

Since a swap is an instrument to hedge interest rate risk, it is natural to ask how it compares with other ways of hedging interest rate risk. We now compare swaps with two alternatives: interest rate futures and debt refinancing.

(a) Swap Versus Interest Rate Futures

What is a futures contract?: *An interest rate futures contract* is an exchange-based contract (as opposed to over-the-counter) to buy or sell a particular financial asset (such as a T-bill) for a specific price at a prespecified date in the future.

Before we can compare swap with a futures contract, you should be aware of how a futures contract can be used to hedge. Consider an S&L with long-term fixed-rate mortgages as assets and short-term CDs as liabilities. Suppose this S&L were to short (sell) a CD futures contract, that is, it could promise to deliver (sell) at a *fixed* price. Then, if interest rates rise in the future, the market value of the CD falls and thus the S&L receives a cash inflow equal to the (positive) difference between the fixed delivery price and the market value of the CD.[27] On the other hand, if interest rates fall and the market value of the CD rises as a result, the S&L will experience a loss. Thus, the gain to the S&L if rates rise is offset by the loss if rates fall. In this way, the S&L's interest rate exposure is hedged.

Advantage of a swap over a futures contract: Interest rate futures are *standardized* contracts with specific delivery dates and specific types of instruments.[28] Thus if you wish to hedge the interest rate risk on a financial claim that is not one of the deliverable instruments on which futures contracts are written, you must choose a futures contract on a deliverable that most closely resembles the claim you wish to hedge. Since the resemblance will be imperfect, you will bear *cross-hedging* risk. Moreover, even if the resemblance were perfect, you would bear *basis risk* (the risk that the relation between the spot and futures prices will change randomly). The major advantage of a swap contract over a futures contract is that a swap can be tailored to suit the customer's need because it is not a standardized contract. Thus, better interest rate hedging is often possible with a swap than with a futures contract. Note, however, that swaps are increasingly becoming more standardized and hence similar to futures contracts, but with longer hedging periods.

Disadvantages of a swap: (i) Imperfect standardization means that it is not always easy to find a counterparty to the desired swap transaction. That is, futures contracts are more liquid than swaps contracts. (ii) Related to (i), the highly customer-specific nature of swaps means that *search costs* may be significant in some transactions. These costs will be passed on to the swapping parties by the swap broker, in the form of a higher fee. Thus, customers face higher transactions costs with swaps than with

27. These are not the S&L's own CDs, but rather a standardized contract. The S&L would not actually buy the CD, but just receive cash settlement.

28. Deliverables in interest rate futures are: T-bills, T-notes, T-bonds, bank and Eurodollar CDs, sterling CDs and gilts, and Ginnie Maes.

futures.[29] (iii) There is a greater risk of nonperformance (default) with a swaps contract than with a futures contract. This is because the exchange guarantees execution with a futures contract, whereas with a swap one party could be left in the cold if the other party reneges and there is no (back-up) guarantee by the swap broker. If there is a back-up guarantee by the swap broker, then the swap broker plays the role of a clearinghouse. But even in this case, there is the possibility of nonperformance by the swap broker.

(b) Swaps Versus Refinancing:
How do you hedge risk by refinancing?: One simple way for a firm to adjust its interest rate exposure is to *directly* refinance. That is, suppose a firm has fixed-rate liabilities and desires floating-rate liabilities. It could simply repurchase its fixed-rate liabilities, financing the repurchase by issuing floating-rate liabilities. Why is this simple approach not always preferred to swaps and futures?

Advantages of a swap over debt refinancing: (i) Swaps avoid many of the transactions costs encountered with debt refinancing, such as legal fees, advertising, and regulatory restrictions. This is because a swap is *not* considered *new* borrowing or a public offering. Rather, it is only regarded as an exchange of interest payments on existing liabilities. (ii) Swaps also avoid many disclosure requirements of new financing because they are not considered new borrowing. This may be of importance to firms that wish to protect the confidentiality of strategic information. (iii) Many firms with low credit ratings pay a higher differential on fixed-rate debt, relative to floating-rate debt, than higher quality firms do. Such low-quality firms may wish to borrow in the floating-rate market and then swap these floating-rate liabilities for fixed-rate liabilities, perhaps avoiding some of the credit risk premium they would need to pay on newly issued debt. Thus, an important reason for the emergence of interest rate swaps (given the availability of direct debt refinancing) may well be that the search costs and credit evaluation costs encountered in nonintermediated (public debt market) transactions can be effectively lowered by financial intermediaries (swap brokers) who specified in mitigating such informational frictions.[30]

Other Contingent Claims: Credit Derivatives

An important development in the contingent-claims markets that banks are involved in is *credit derivatives*, a market that barely existed until 1997, but is now trillions of dollars in magnitude. The basic idea behind a credit derivative is simple. A lender essentially purchases from a third party a put option on the borrower's debt, which entitles the lender to "put" the debt, if its value is impaired due to, say, default, to the third party. This way the lender purchases insurance against credit risk. Banks have been active players on both sides of this market, both as purchasers of credit risk insurance and as sellers of this insurance.

Figure 8.8 shows the explosive growth of the credit derivatives market from 1997 through 2006.[31] Second, because of the spread of securitization to the credit-derivatives market, there is *pooling* and tranching of diverse credit risks. This

29. Swap brokers charge on average an arrangements fee of about 25 basis points, not including an additional fee for guaranteeing the contract (that is, the additional fee for acting as an asset transformer).

30. See Campbell and Kracaw (1991) for an analysis of swaps along these lines.

31. See The Economist, August 20–26, 2005.

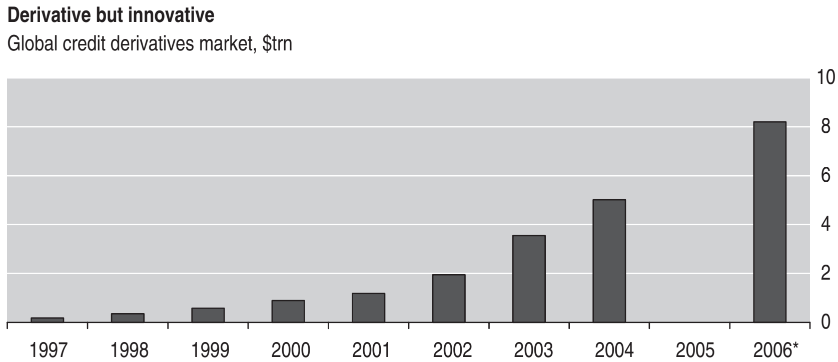


FIGURE 8.8 Growth in Credit-Derivatives Market

* Forecast

Source: British Bankers' Association.

enables idiosyncratic shocks to individual credit risks to be diversified away and risks to be spread out over many market participants and hence managed more adroitly.

The development of the credit derivatives market has been facilitated by the growing standardization of credit-derivatives contracts, and the creation of indices that offer hedges against pools of U.S. as well as non-U.S. corporate credits, such as European and Japanese corporate credits. While initial credit derivatives were simple credit default swaps involving single companies, much of the recent growth has been via pooling together of numerous credits and then tranching as in other forms of securitization (see the next chapter). Securitization has also invited significant institutional participation in this market. It is estimated that a large percentage of the trading volume in credit derivatives is accounted for by hedge funds (discussed earlier in the book).

Risks for Banks in Contingent Claims

An Overview of Risks

With the enormous growth in the contingent claims products offered by banks, there has been growing concern that their balance sheets grossly underestimate their risks. The reasons for concern are twofold. First, because contingent claims have not required reserves or capital to support them, it has been quite tempting for banks to sell these claims in large volume, so that the off-balance sheet risk for any individual bank can become substantial. If the bank is lucky, it earns its fee revenues from sales of these claims, without suffering the adverse consequences of risk. But if things go sour, the bank could experience capital impairment, which in turn could provide further incentives for it to take risk because of the (put option) nature of deposit insurance.[32] Of course, under the BIS capital guidelines, banks are now

32. More on this in later chapters.

required to hold capital against OBS claims, so that their attractiveness to banks may decline somewhat. However, some OBS claims are exempted from capital requirements, for example, loan commitments with maturities less than one year. Second, contingent claims often create interlocking relationships across banks that could strengthen the *contingent effect* of bank failures.[33]

We can now address the risks in individual contingent claims.

Risks in Loan Commitments

Regulators regard loan commitments as the second-riskiest contingent claim, just behind standby L/Cs, which are discussed below. A bank faces three types of risks in loan commitments: (i) the risk that it may have to lend at a lower-than-spot-market margin or even a negative margin due to fixity in the commitment rate, (ii) the risk that it may be forced to lend to higher-risk customers than its spot market, given the pool, that is, borrowers it would not have loaned to in the spot market, given the conditions at the time of commitment takedown, and (iii) the risk that it may have to fund commitment when its own liquidity is low and costly to replenish. Each of these risks is discussed below.

(i) The Risk of Lending at Low Margins: One risk in loan commitments is that the bank may be compelled to grant loans at interest rates that either reduce profits relative to spot lending opportunities, or result in direct losses. This risk is lower per dollar of commitment with variable-rate commitments, but is present nonetheless. For example, the commitment may permit the customer to borrow at prime plus 1 percent. The bank cannot be assured, however, that the customer's creditworthiness will not deteriorate between the time the commitment is issued and the time that it is exercised. A customer who is "prime plus one" when the prime is 10 percent is likely to be riskier than "prime plus one" when the prime climbs to 20 percent. The creditworthiness of a borrower can be expected to vary inversely with market interest rates since higher interest rates will usually absorb a greater fraction of the borrowing firm's cash flows. Thus, even under a variable-rate commitment the bank is exposed to the risk of earning a lower interest rate on commitment loans than it would if the same funds were invested in spot loans with the same credit risk.

The prime-times contract addresses this linkage between the prime and the customer's add-on. This contract imposes opportunity costs of the type described above only when the customer's appropriate add-on rises by more than the percent indicated in the commitment. Although the prime-times contract imposes some risk on the bank, the bank's exposure per dollar under this contract is less than for the prime-plus contract. This is because customer add-ons for spot borrowing tend to increase exponentially (as in 2, 4, 8, 16 . . .) with increase in the prime rate, rather than proportionally (as in 2, 4, 8 . . .) as in the prime-times commitment contract.

Even if the customer's creditworthiness does not vary with market interest rates, the bank faces a risk with loan commitments due to sluggishness in the prime rate.[34] This sluggishness means that the bank's funding cost is only imperfectly correlated with the prime, so that as the bank's funding cost changes with movements in market

33. See Andrews and Sender (1986).
34. Recall the discussion in Chapter 6 of the sluggishness in the prime relative to market interest rates.

interest rates, the bank will need to adjust the add-on (or multiple) to the prime that it charges the borrower. In a variable-rate commitment this add-on (multiple) is held fixed, so that as interest rates rise, the add-on (multiple) the bank should charge to break even grows larger and larger than the commitment add-on (multiple). At sufficiently high prime rates, the bank's spread between the commitment rate and its cost of funds may well invert and become negative. Of course, the reverse is true when rates are falling, but there is an asymmetry due to the option nature of the commitment since the customer will simply let the commitment expire unexercised. This is a risk that the bank does not face with spot lending because it can always adjust the add-on to the prime to reflect the prime's sluggishness in responding to market interest rate movements.

(ii) The Risk of Being Forced to Lend to Excessively Risky Customers: Loan commitments also may expose banks indirectly to increased *credit risk*. The relationship between interest rate and credit risks is manifested in two ways. First, as interest rates increase and become more volatile, the economic value of the cash flows generated by the customer's investments may become smaller and more uncertain. That is, when inflation increases, the percentage spread between nominal and real interest rates is likely to widen more than the percentage spread between nominal and real cash flows. Thus, under both fixed- and variable-rate commitments, high and volatile interest rates can expose the bank to increased credit risk. Second, high and volatile interest rates can increase credit risk through an asset substitution effect (recall the credit rationing discussion in Chapter 6) that is more likely with variable-rate commitments. The customer can be expected to adapt to a higher borrowing rate by choosing investment projects with higher expected payoffs, and these usually would have been rationed in the spot market. Note that this risk is different from that discussed under (i) in that the risk there is that the bank's profit margin on "acceptable" borrowers—those it would not have rationed in the spot market—may become too low, and the risk here is that the bank may have to lend to "unacceptable" borrowers. This risk is obviously absent in spot lending.

Of course, the MAC clause is supposed to enable the bank to extricate itself from a commitment to a borrower whose financial condition has deteriorated significantly. The safety provided by this clause may be limited, however, due to the bank's reputation-driven reluctance to invoke the MAC.[35]

(iii) The Risk of Funding Commitments in Low-Liquidity Periods: There are two reasons why a bank may find itself liquidity constrained. One is that there may be a marketwide decline in liquidity. The other is that there may be bank-specific problems that cause familiar sources of liquidity to become substantially more expensive or even dry up. In either case, commitments become costlier to fund, a risk that is not encountered with spot lending.

Risks in Letters of Credit

Commercial L/Cs are used for routine trade transactions and carry with them credit risk, whereas standby L/Cs are mainly financial guarantees under which, in exchange for fees, banks guarantee a variety of financial obligations of borrowers to specified

35. See Boot, Greenbaum, and Thakor (1993).

third parties. These pledges include credit enhancement facilities to municipal borrowers, commercial paper issuers, and those involved in securitizations.[36] Banks consider these guarantees risky because they are irrevocable and are activated by borrower financial distress.

There are three basic types of risks faced by banks in L/Cs: (i) credit risk, (ii) documentation risk, and (iii) political risk. We discuss each in turn.

(i) Credit Risk: Commercial and standby L/Cs differ in that the bank pays on performance with commercial L/Cs and on nonperformance with standby L/Cs. This difference is not that significant for the bank's risk exposure, however, since it is not the ability of the debtor to perform the stipulated task that determines the bank's risk. Rather, the bank's risk in both cases turns on the debtor's reimbursement of the bank. Thus, one risk the bank faces with either a commercial or a standby L/C is *routine credit risk*. Bankers have recognized the similarity between the risks faced in normal lending and in issuing L/Cs. Although L/Cs are originated in a number of "nontraditional" divisions within the bank, such as municipal or corporate finance divisions, bankers say they apply the same credit screening procedures to standby L/Cs that they apply to their loans.

It is often claimed that standby L/Cs are more risky than commercial L/Cs. One reason for this claim is that standby L/Cs pay on nonperformance, whereas commercial L/Cs pay on performance. As noted above, this distinction is not that significant for assessing the bank's risk exposure across the two L/Cs. Another reason why standbys are considered riskier than commercial L/Cs is that the latter routinely generate collateral in the form of goods in storage or transit (". . . commercial L/Cs are self-liquidating"), whereas standbys may be unsecured. However, collateral does not always accompany commercial L/Cs, and standby L/Cs are not always unsecured. Furthermore, financial distress often accompanies a decline in the value of the customer's collateral, so that collateral may offer the bank only limited protection in the case of L/Cs.

It is true, nonetheless, that regulators and banks consider standby L/Cs as the riskiest of all the contingent claims offered by banks. One reason may be that standby L/Cs are used to cover almost any contingency, whereas commercial L/Cs are used for routine trade transactions. Thus, standby L/Cs may be riskier simply because they cover a variety of contingencies.

(ii) Documentation Risk: Documentation presents another source of risk in commercial L/Cs. Although this risk is routinely accepted by banks, a Federal Reserve survey found that in approximately 35 percent of the cases examined, documentation *failed* to conform to the requirements of the L/C. Improper documentation can invalidate a contract and prompt the buyer to refuse to accept delivery. In this case, the bank will be forced to find a buyer on its own, or to take possession of the goods.

(iii) Political Risk: U.S. exporters are sometimes unfamiliar with the foreign bank issuing an L/C. They may also be concerned about the political climate in the importer's country. The exporter may, in these cases, obtain a confirmation from a U.S. bank that is then obliged to make payment if the drawee is unable to do so. The confirming (American) bank faces two risks. One is that the issuing (foreign) bank will default, and the other is the political risk of exchange controls.

Risks in Interest Rate Swaps

Although swaps aggregate to trillions of dollars, this figure is the sum of the principal amounts involved in the deals. In fact, only the interest rate streams are at risk since each issuer retains its obligation for its own principal. Moreover, as swap brokers, the liabilities of banks are quite limited. There are two types of risks in swaps: (i) counterparty risk and (ii) legal risks. We discuss each now.

(i) Counterparty Risk: The biggest risk for the bank is that one of the swap partners will be unable to make its interest payments. The bank then has to either assume the interest payments for the defaulting party or replace the defaulting party; this is essentially an exposure to interest rate risk. As an overall assessment, however, swaps appear to be the least risky of the three major contingent claims that we have discussed.

(ii) Legal Risks: There may be significant *hidden* legal risks in swaps that have only recently begun to surface. For example, there have been cases in which a failed bank launched court proceedings against a solvent counterparty bank over a swap, claiming that the counterparty should have honored the swap contract even after the failed bank declared bankruptcy. This was done despite the fact that the terms of the contract allowed for "limited two-way payments," under which if one party defaulted, the other was *not* liable for any payments under the contracts. (These are in contrast to "full two-way payment" contracts, under which both parties are obligated to make full payments under the swap contract even if one party defaults on other obligations). Solvent counterparty banks often make good on their liabilities (despite *no* contractual obligation to do so) because of reputational concerns and nervousness about whether their lack of contractual obligations would hold up in court if they refused to perform. This event vividly illustrates the manner in which a bank can use the contractual discretion in a contingent claim to (optimally) write down its reputational capital in order to conserve financial capital. It also shows that this trade-off is bank specific, since different banks have different reputations and different levels of financial capital.

Regulatory Issues

The *Basle Accord* (Basel 1) reached under the auspices of the *Bank for International Settlements* (BIS) in 1987, stipulated a new set of *capital guidelines* under which loan commitments with maturities under 1 year are not subject to capital requirements, whereas longer-maturity commitments have a 4 percent capital requirement (which is half the capital requirement against most loans). Moreover, a commitment that the bank can unconditionally cancel without cause and for which it conducts an annual credit review (to decide whether it should be continued) will be regarded as having a maturity under 1 year. Standby L/Cs or other types of bank guarantees are also subject to capital requirements. The capital requirement against standby L/Cs is 8 percent.

OBS items continue to be free of cash-asset *reserve requirements*. Thus, a bank need not hold cash-asset reserves against a loan commitment until the customer exercises it, at which stage the amount taken down is a loan. If the bank funds this loan with deposits, then it must hold reserves against these deposits. But if the bank chooses to sell the loan or securitize it (see Chapter 9), it can avoid the reserve requirement.

The *accounting treatment* of contingent claims is another issue. Although many contingent claims impose contingent liabilities on banks, these liabilities do not appear on the balance sheet, except as footnotes. On the other hand, the fee collected by the bank is recognized on the income statement, albeit on the basis of an amortization schedule that requires recognition over the life of the contingency. Since the cash generated by the fee income augments the book value of the bank's assets, whereas the (offsetting) contingent liability does not increase the book value of the bank's liabilities, the sale of contingent claims permits a bank to artificially inflate the book value of its net worth. Moreover, since the fees collected and the contingent liabilities imposed can be expected to be larger during periods of greater interest rate volatility, the inflation of book net worth will be greater when interest rates are more volatile. This would not be the case if the liability diminishes at the pace the income is recognized, but this seems unlikely.

Conclusion

In this chapter we have reviewed the theory of commercial bank contingent claims and have commented upon their magnitude and growth. Loan commitments and L/Cs are an outgrowth of commercial bank lending in much the same way that agricultural futures markets are an outgrowth of grain trade. By the late 19th century, commercial banks had adopted the practice of informally assuring renewal of the short-term notes of their customers. It was a short step from such agreements to more formalized commitments. The emergence of loan commitments of the types observed today can be traced back to the early 1920s.[36] That period marked a shift in attitude within the banking community from the "real bills" doctrine,[37] focusing on short-term self-liquidating commercial loans, to the "shiftability" theory of funds management. The latter finds liquidity in a wider variety of bank claims, providing the basis for an increased willingness by bankers to precommit loans. Forward lending quickly developed into an integral part of commercial banking.

The emergence of liability management in the 1960s along with the tight credit conditions of 1966 and 1969 increased loan commitment activity. Tight credit conditions induced borrowers to seek more loan commitment and the advent of liability management provided banks with new means of raising the funds required to meet this demand. The late 1960s and 1970s were characterized by interest rates that were both higher and more volatile. Increasing inflation led to greater loan demand and periodic credit crunches increased the demand for credit lines. Banks became less willing to offer fixed-rate commitments in the face of highly unpredictable

36. See Wood (1983).

37. The main point of the "real bills" doctrine was the idea that a sufficient condition for desirable monetary policy is that all banks, including the central bank, restrict their lending to "nonspeculative" loans secured by "real" collateral, that is, inventories and other tangible assets. The legislation that established the Federal Reserve System was influenced by this doctrine. A criticism of this doctrine is that it leads to a procyclical monetary policy since the Federal Reserve makes more credit available to banks in "good times" when they have sufficient eligible collateral and less in "bad times" when they have fewer assets to serve as eligible collateral. Consequently, monetary policy exaggerates and exacerbates the business cycle.

interest rates and thus, many began to "float" the prime (the prime rate changed 40 times in 1980 as opposed to 23 times in the 13 years from August 1955 to December 1968) and offer variable-rate commitments that provided little or no protection against changes in the prime.[38] Moreover, the increased interest rate volatility was accompanied by elevated volatility in the capital market and the foreign exchange market. This made risk management critical for the customers of banks, and banks provided this service through a host of new derivatives and other contingent claims.

We have also discussed how recent changes in regulation have led to the imposition of capital requirements on contingent claims. These regulatory changes mean that the supply-side incentives for commercial bank contingent claims have been weakened somewhat. Despite this, we expect contingent claims to grow in importance in the future. As the banking industry continues to be reconfigured, it will be interesting to examine how the market for contingent claims is divided between commercial banks and their newer, nonbank competitors.[39]

Case Study Youngstown Bank

Introduction

John Standard has been the chief executive officer (CEO) of Youngstown Bank since the summer of 1998. Before taking this position, he had been a vice president of operations for Interbank, a large regional bank. One of the primary reasons that he was hired by Youngstown Bank was his experience with a large operating department. At the time, Youngstown Bank had been going through some difficulties related to inefficient operating procedures, and Mr. Standard had acquired a reputation at Interbank for strong motivational and organizational skills. His management of Youngstown has been almost flawless, and the institutional culture of the bank takes great pride in the fact that the bank is a very "tight ship."

Youngstown Bank has been in business in Youngstown, Arizona, since 1910. When John Standard was brought in as CEO in 1998, the stock price was at 4½, down from a high of 10. The previous CEO was the son of the founder, and he had resisted the replacement of legacy systems with more modern information processing infrastructure, allowing the operating departments to languish in mediocrity. Prior to Mr. Standard's arrival, people barely even knew what the bank's policies were on loans! The only kinds of products Youngstown Bank offered were simple fixed-rate loans. John Standard changed all that. He put together a set of standard procedures for loans and loan commitments, and attempted to tailor the bank's policies to the risk and liquidity needs of its customers. And the stock price responded; by the end of 1999, Youngstown Bank's stock price had doubled to $9, and continued to rise through 2000.

38. See Arak, Englander, and Tang (1983).

39. Investment banks, for example, have been extremely active in innovation of contingent claims. There has been a veritable explosion in highly customized options that trade in the over-the-counter market. One advertisement promised to supply markets in "min-max-zeros, range forwards, cylinder options, reverse forward options, quantos, zero cost collars, compound options, targets, scouts, flying hedges, moon rockets, the almost impossible to understand option," and so on, and ended with the promise, "We'll write it. You name it." See Rubinstein and Reiner (1992).

But starting in 2001, the bank's stock price has been languishing. Even though the bank's basic structure has not changed and profitability is good, the stock price has simply not moved upward over time, although the stock prices of some competing banks have moved up significantly. The major shareholders in the bank aren't too upset yet, but there have been a few grumblings. Standard realizes that there could be major trouble down the line unless he can find a way to get the share price up. He decides to call in his chief financial officer (CFO), Bryan Shelton, to discuss the stock price situation.

The Initial Meeting

Standard: Come on in, Bryan, and have a seat. Let's get right down to business here. I'm worried about our stock price performance lately. You've been with Youngstown Bank for three years now—what was the stock price when you got here?

Shelton: It was right around 37, I think.

Standard: Well, it is just over 40 now. We closed at 40 ¼ yesterday. That's only 3 dollars in 3 years! What is going on? I don't understand it. Why is our stock price so low? Take a look at how our market-to-book ratio compares with that of our competitors. It is in the dirt! (See Exhibit A). Why?

Shelton: That's a good question. Considering how precisely we control everything, and considering that our profits and cash flows are still looking good, I don't know of any reason why the stock should be down. I'm tempted to just say that the market is failing to recognize our value. Maybe they'll come around when we post good numbers again next quarter.

Standard: Well, you might be right, but I'm uncomfortable. Maybe the market is reacting to something that we don't know about. I think we should look into this some more, and try to get to the bottom of it.

[*The meeting ends on that note, and Mr. Shelton says that he will look into the matter carefully and report back. He agrees that they should meet a week later to discuss the issue again.*]

The Second Meeting

Shelton: Well, I've looked into this some more, and frankly I'm still puzzled. Take a look at these numbers. Our current balance sheet looks good, and compares very favorably with the way it looked during 2000, the heyday of our stock price rise (see Exhibit B). Our key rations look just fine, too, compared to 2000 (see Exhibit C). Moreover, we also seem to be doing well relative to industry averages (see Exhibit D).

Standard: This all looks great, just like I thought it would. Look at this one. (*He points at Exhibit D.*) Our return on assets is great. So what do you think?

Shelton: Well, one of the people I had helping me to put these numbers together for you suggested that we might want to think about our loan commitments, which don't appear on our balance sheet. Maybe those are dragging our stock price down.

Standard: That doesn't make sense. Our policies on loan commitments haven't changed, have they? What kind of data do you have on those?

Shelton: Well, take a look at these. (*He pulls out Exhibits E and F.*) These show the history of interest rates and the fees that we charge for loan commitments. I checked on the kinds of borrowers who've been buying these commitments, and the quality of the borrowers seems to be in line with our history. To tell you the truth, I'm still struggling with what all this stuff means. I don't see that anything has changed anywhere. But our stock price . . .

Standard: Well, all I can tell you is keep working on it. See if you can find anything here that will help explain why our stock price is low. Is there something that we've overlooked? Is the bank in some danger that we've failed to realize?

[*Again, the meeting ends and they agree to meet in a week. This time, Standard has some specific questions to which he wants answers. Shelton plans to go over everything carefully, looking for some explanation for the poor performance of the stock price, an explanation that takes into account all the facts about the bank's situation.*]

The Numbers

Exhibit A
YOUNGSTOWN BANK, INC.
Market-to-Book Ratio
Comparison to Industry

Year	Youngstown	BancFirst	Industry
1991	.51	1.21	1.18
1992	1.00	1.11	1.08
1993	1.43	1.23	1.13
1994	1.47	1.32	1.21
1995	1.60	1.43	1.31
1996	2.13	1.87	1.53
1997	1.35	1.41	1.41
1998	1.18	1.11	1.20
1999	1.35	1.32	1.27
2000	1.41	1.31	1.34
2001	1.21	1.40	1.47
2002	.95	1.65	1.53
2003	.81	1.89	1.66
2004	.78	1.86	1.63

Exhibit B
YOUNGSTOWN BANK, INC.
Year-End Balance Sheets (in Thousands of Dollars)

	2000	2005
Assets		
Cash & Due	125,000	129,000
Marketable Securities	200,000	400,000
Loans:		
Real Estate	190,000	385,000
Commercial and Industrial	315,500	744,000
Consumer	140,500	153,742
All Other	131,400	142,300
Less Unearned Income:		
Allowances for Possible Loan Losses	1,316	1,500
Total Loans	776,084	1,423,542
Other Assets	78,000	150,000
Total Assets	1,179,084	2,102,542
Liabilities and Equity		
Liabilities:		
Deposits	1,000,020	1,775,420
Federal Funds Purchased	75,000	102,000
Other Liabilities	63,000	90,000
Total Liabilities	1,138,020	1,967,420
Equity Capital:		
Preferred and Common Stock	11,000	35,122
Surplus	14,064	42,000
Undivided Profits and Reserves	16,000	58,000
Total Equity Capital	41,064	135,122
Total Liabilities and Equity	1,179,084	2,102,542

Note: Volume of outstanding loan commitments in 2000 was $1,000,500 and 2005 was $4,320,000.

Exhibit C
YOUNGSTOWN BANK, INC.
Comparison of Performance for 2000 and 2005

	2000	2005
Net Income (in thousands of dollars)	8,607	16,820
Return on Assets (in percentage)	0.73	0.80
Total Liabilities to Total Assets	0.97	0.94
Total Liabilities to Common Equity	27.71	14.56

Exhibit D
Various Industry Ratios for 2005
(Averages for Similarly Sized Banks)

	Youngstown	Average
Return on Assets	0.8	0.6
Total Liabilities to Total Assets	0.94	0.97
Total Liabilities to Common Equity	14.56	21.3

Exhibit E
Interest Rate History
(Annualized Interest Rates in Percentage)

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1991	7.95	8	8	8	8.27	8.63	9	9.01	9.41	9.94	10.94	11.55
1992	11.75	11.75	11.75	11.75	11.75	11.65	11.54	11.91	12.9	14.39	14.55	15.3
1993	15.25	15.63	18.31	17.77	15.57	12.63	11.48	11.69	12.23	14.79	16.06	17.1
1994	20.16	19.43	18.05	17.15	19.61	20.03	20.39	20.5	20.06	18.45	16.84	16.75
1995	15.75	16.56	16.5	16.5	15.5	15.5	14.26	14.39	13.5	12.52	11.85	11.5
1996	11.16	10.98	10.5	10.5	10.5	10.5	10.5	10.89	11	11	11	11
1997	11	11	11.21	11.93	12.39	12.6	13	13	12.97	12.58	11.77	11.06
1998	10.61	10.5	10.5	10.5	10.31	9.78	9.5	9.5	9.5	9.5	9.5	9.5
1999	9.5	9.5	9.1	8.83	8.5	8.5	8.16	7.9	7.5	7.5	7.5	7.5
2000	7.5	7.5	7.5	7.75	8.14	8.25	8.25	8.25	8.7	9.07	8.78	8.75
2001	8.75	8.51	8.5	8.5	8.84	9	9.29	9.84	10	10	10.05	10.5
2002	9.8	9.1	8.2	7.8	7.2	6.3	5.32	5.01	7.73	5.21	5.09	8.3
2003	9.2	8.3	7.4	7.1	6.2	5.5	5.1	4.8	4.5	6.2	9.1	8.1
2004	6.1	3	3	3	4	6.83	9.23	9.3	10.2	8.5	7.43	8.91

Exhibit F
Loan Commitment Prices
(Average in Basis Points)

	Commitment Fee	Annual Servicing Fee	Usage Fee
1994	12.5	12.5	25.0
1995	12.0	12.0	25.0
1996	12.0	12.0	25.0
1997	12.5	12.0	22.5
1998	12.5	12.5	22.5
1999	12.5	12.5	21.5
2000	12.5	12.5	22.5
2001	12.5	12.5	25.0
2002	12.0	12.5	25.0
2003	12.5	12.5	25.0
2004	14.0	12.5	27.5

The Assignment

Mr. Standard gives Mr. Shelton these specific questions:

1. Is the lack of upward movement in the stock price evidence of market irrationality or overreaction, or is something else going on?
2. What should the bank do? What strategies should the bank pursue? What, if any, are the major dangers faced by the bank?

Review Questions

1. What is an off-balance sheet contingent claim, and what are the major types of contingent claims observed today?
2. Define a loan commitment and briefly discuss the different types of loan commitments.
3. Provide discussion of the supply-and-demand-side motivations for loan commitments.
4. It has been claimed that a bank loan commitment has an isomorphic correspondence with a common stock put option. How valid is this claim?
5. Discuss a commercial L/C, a standby L/C, and a bankers acceptance.
6. What is an interest rate swap and how does it work?
7. What is the role of a swap broker in an interest rate swap transaction?
8. Discuss three variations of the "plain vanilla" swap.
9. What are swaptions, caps, collars, and floors?
10. What are the advantages and disadvantages of an interest rate swap relative to a futures contract as a hedging instrument?
11. What is the advantage of a swap over direct financing for hedging interest rate risk?
12. Discuss the risks faced by commercial banks in loan commitments, letters of credit, and interest rate swaps.
13. Suppose a borrower knows at $t = 0$ that it will have available at $t = 1$ an opportunity to invest \$175 in a risky project that will pay off at $t = 2$. The borrower knows that it will be able to invest in one of two mutually exclusive projects, S or R, each requiring a \$175 investment. If the borrower invests in S at $t = 1$, the project will yield a gross payoff of \$310 with probability 0.8 and zero with probability 0.2 at $t = 2$. If the borrower invests in R at $t = 1$, the project will yield a gross payoff of \$330 with probability 0.6 and zero with probability 0.4 at $t = 2$. The borrower's project choice is not observable to the bank.

    The riskless, single-period interest rate at $t = 0$ is 12 percent. It is not known at $t = 0$ what the riskless, single-period interest rate at $t = 1$ will be, but it is common knowledge that this rate will be 8 percent (with probability 0.6) or 15 percent (with probability 0.4). Assume universal risk neutrality and that the borrower has no assets other than the project on which you (as the lender) can have any claim.

    Suppose you are this borrower's banker and both you and the borrower recognize that this borrower has two choices: (i) it can either do nothing at $t = 0$ and simply plan to borrow in the spot market at the interest rate prevailing for it at $t = 1$, or (ii) it can negotiate at $t = 0$ with you (or some other bank) for a loan commitment that will permit it to borrow at predetermined terms at $t = 1$. What advice should you give this borrower? Assume a competitive loan market in which each bank is constrained to earn zero expected profit.

14. The following is an excerpt from "A Friendly Conversation." Critique it.

    Appleton: That's simple, Mike. The BIS stipulations are *minimum* levels, whereas the Treasury proposal gives banks choices above the BIS minima. What bothers me about the BIS guidelines, though, is that they also require banks to hold capital against *off-balance sheet items*. When these items get on the balance sheet, there is another capital requirement against them, so aren't we in a sense double counting?

Butterworth: Not really, because there is not simultaneity involved. I think that with a trillion dollars in outstanding loan commitments alone, the issue of the contingent liability exposure of American banks is something that we just have to come to grips with. The way that RAP (Regulatory Accounting Principles) and GAAP (Generally Accepted Accounting Principles) have dealt with these contingent liabilities has been deplorable. I strongly believe depository institutions should be made to recognize these liabilities *on* their balance sheets, not merely in footnotes.

Appleton: Beth, I think you are getting a bit carried away. Nobody has any idea how these contingent liabilities should be valued, so how do you quantify your exposure?

Butterworth: Speak for yourself, Alex. There *are* valuation models available, although I will admit they are far from perfect. But even noisy information is better than none.

15. Critique the following excerpt from "A Friendly Conversation."

Moderator: Hold it there people. Remember, I cannot be here forever. I thought we were discussing banking reform and deposit insurance. Does all this talk about off-balance sheet activities have anything to do with deposit insurance?

Butterworth: That is a good question, Mike. I honestly do not know, but my guess is that contingent liabilities represent a hidden liability for the deposit insurance fund. The more contingent liabilities the bank has, the more risk there is in the banking system.

Appleton: As both of you know, I believe that off-balance sheet activities are the future of banking, so Beth's views on this trouble me. Perhaps she has some evidence to support her claim?

Butterworth: No, Alex I do not. But I will research the matter.

References

Antl, Boris (ed.), *Swap Financing Techniques*, London: Euromoney Publications Limited, 1983.

Arak, M.A., A.S. Englander, and E.M.P. Tang, "Credit Cycles and the Pricing of the Prime Rate," *Federal Reserve Bank of New York Quarterly Review*, Summer 1983.

Avery, Robert B., and Allen N. Berger, "Loan Commitments and Bank Risk Exposure," *Journal of Banking and Finance* 15, 1991, 173–192.

Baecher, Eileen, "Swaps and the Derivative Markets," manuscript, Fixed Income Research, Goldman Sachs & Company, January 1991.

Beidleman, Carl R., *Financial Swaps: New Strategies in Currency and Coupon Risk Management*, Homewood, Illinois: Dow Jones-Irwin, 1985.

Berger, Allen N., and Gregory F. Udell, "Some Evidence on the Empirical Significance of Credit Rationing," working paper, Board of Governors of the Federal Reserve System, December 1990.

Berkovitch, Elazar, and Stuart I. Greenbaum, "The Loan Commitment as an Optimal Financing Contract," *Journal of Quantitative Analysis* 26, 1990, 83–95.

Boot, Arnoud, and Anjan V. Thakor, "Off-Balance Sheet Liabilities, Deposit Insurance and Capital Regulation," *Journal of Banking and Finance*, 15, 1991, 825–846.

Boot, Arnoud W.A., Stuart I. Greenbaum, and Anjan V. Thakor, "Reputation and Discretion in Financial Contracting," *American Economic Review* 83–5, December 1993, 1165–1183.

Boot, Arnoud, Anjan V. Thakor, and Gregory F. Udell, "Credible Commitments, Contract Enforcement Problems and Banks: Intermediation as Credibility Assurance," *Journal of Banking and Finance* 15, June 1991, 605–632.

———, "Competition, Risk Neutrality and Loan Commitments," *Journal of Banking and Finance* 11, September 1987, 449–471.

Campbell, Tim S., "A Model of the Market for Lines of Credit," *Journal of Finance* 33, March 1978, 231–244.

Campbell, Tim S., and William Kracaw, "Intermediate and the Market for Interest Rate Swaps," *Journal of Financial Intermediation* 1–4, December 1991, 362–384.

Deshmukh, Sudhakar D., Stuart I. Greenbaum, and George Kanatas, "Bank Forward Lending in Alternative Funding Environments," *Journal of Finance* 37, September 1982, 925–940.

Duca, John, and David D. Vanhoose, "Loan Commitments and Optimal Monetary Policy," *Journal of Money, Credit and Banking* 22–2, May 1990, 178–194.

Glick, Reuven, and Steven E. Plaut, "Money Demand and Off-Balance Sheet Liquidity: Empirical Analysis and Implications for Monetary Policy," *Journal of Accounting, Auditing and Finance* 4, 1989, 147–159.

Greenbaum, Stuart I., George Kanatas, and Itzhak Venezia, "Loan Commitments and the Management of Uncertain Demand," *Journal of Estate Finance and Economics* 4, December 1991, 351–366.

Greenbaum, Stuart I., and Itzhak Venezia, "Optimal Exercise of Loan Commitments Under Adaptive Pricing," *Journal of Financial Research* 8–4, Winter 1985, 251–263.

Greenbaum, Stuart I., John Soss, and Anjan V. Thakor, "Understanding Commercial Bank Contingent Liabilities," monograph, *Bank Administration Institute*, 1986.

Hawkins, Greg, "An Analysis of Revolving Credit Agreements," *Journal of Financial Economics* 10, 1982, 59–81.

Holmstrom, Bengt, and Jean Tirole, "Market Liquidity and Performance Monitoring," *Journal of Political Economy*, 1993.

"International Banking: A Comedy of Errors," *The Economist*, April 10–16, 1993.

James, Christopher, "An Analysis of Bank Loan Rate Indexation," *Journal of Finance* 37, 1982, 809–825.

———, "Self-Selection and the Pricing of Bank Services: An Analysis of the Market for Loan Commitments and the Role of Compensating Balance Requirements," *Journal of Financial and Quantitative Analysis* 16, December 1981, 725–746.

Kanatas, George, "Commercial Paper, Bank Reserve Requirements, and the Informational Role of Loan Commitments," *Journal of Banking and Finance* 11, September 1987, 425–448.

Kareken, John H., "The Emergence and Regulation of Contingent Commitment Banking," *Journal of Banking and Finance* 11, 1987, 359–377.

Lipin, Steven, "Banks Fear Corporations Will Tap Lines of Credit," *American Banker*, November 7, 1990.

Loeys, Jan G., "Interest Rate Swaps: A New Tool for Managing Risk," *Business Review*, Federal Reserve Bank of Philadelphia, May/June 1985, 17–25.

Maksimovic, Vojislav, "Product Market Imperfections and Loan Commitments," *Journal of Finance* 45, December 1990, 1641–1654.

Melnik, Arie, and Steven E. Plaut, "The Economics of Loan Commitment Contracts: Credit Pricing and Utilization," *Journal of Banking and Finance* 10, June 1986, 267–280.

Melton, W.C., and J.M. Mahr, "Bankers Acceptances," *Quarterly Review*, Federal Reserve Bank of New York, Summer 1981.

Morgan, D., "Bank Credit Commitments and Credit Rationing," working paper, Federal Reserve Bank of Kansas City, Kansas City, Kansas, March 1989.

Rubinstein, Mark, and Eric Reiner, "Exotic Options," manuscript, Leland O'Brien Rubinstein Associates, May 3, 1992.

Shockley, Richard, and Anjan V. Thakor, "Bank Loan Commitments: Data, Theory and Tests," *Journal of Money, Credit and Banking* 29–4, November 1997, 517–534.

Shockley, Richard, "Bank Loan Commitments and Corporate Leverage," *Journal of Financial Intermediation* 4, 1995, 273–301.

Sofianos, George, Paul Wachtel, and Arie Melnik, "Loan Commitments and Monetary Policy," *Journal of Banking and Finance* 14–4, 1990, 677–689.

Thakor, Anjan V., "Do Loan Commitments Cause Overlending?" *Journal of Money, Credit and Banking*, 37–6, December 2005, 1067–1100.

Thakor, Anjan V., "Competitive Equilibria with Type Convergence in an Asymmetrically Informed Market," *Review of Financial Studies* 2, Summer 1989, 49–71.

———, "Commercial Bank Contingent Claims Products," monograph, *The Herbert V. Prochnow Foundation of the Graduate School of Banking*, Madison, Wisconsin, 1988.

———, "Toward a Theory of Bank Loan Commitments," *Journal of Banking and Finance* 6, March 1982, 55–83.

Thakor, Anjan V., and Gregory F. Udell, "An Economic Rationale for the Pricing Structure of Bank Loan Commitments," *Journal of Banking and Finance* 11, 1987, 271–289.

Thakor, Anjan V., Hai Hong, and Stuart I. Greenbaum, "Bank Loan Commitments and Interest Rate Volatility," *Journal of Banking and Finance* 5, December 1981, 497–510.

Wojnilower, A.M., "The Central Role of Credit Crunches in Recent Financial History," Brookings Papers on Economic Activity, 1980, 277–326.

Wolkowitz, B., Peter Lloyd-Davies, B.C. Gendreau, Gerald A. Hanweck, and Michael A. Goldberg, *Below the Bottom Line*: *The Use of Contingencies and Commitments by Commercial Banks*, Staff Study 113, Board of Governors of the Federal Reserve System, Washington, D.C., January 1982.

Wood, John W., "Familiar Developments in Bank Loan Markets," *Economic Review*, Federal Reserve Bank of Dallas, November 1983.

CHAPTER ♦ 9

Securitization

"Robert M. Greer is apartment hunting, even though he doesn't need a place to live. What the Lones Lang Wooton managing director is seeking is the best apartment buildings for inclusion in a securitized mortgage portfolio."

American Banker, October 2, 1990

Glossary of Terms

GNMA: Government National Mortgage Association (see Chapter 5).

FNMA: Federal National Mortgage Association (see Chapter 5).

FHLMC: Federal Home Loan Mortgage Corporation or "Freddie Mac" (see Chapter 5).

FHA: Federal Housing Administration is a federal agency within the HUD Department. The FHA makes no loans, but it operates a variety of loan insurance and subsidy programs designed for low-income housing to help stabilize that segment of the home mortgage market.

Implicit Contract: A term used in economics to designate an implicit understanding between parties about future behavior. There is *no* explicit contract, nor is the promise necessarily legally binding.

GMAC: General Motors Acceptance Corporation is a finance company that is a subsidiary of General Motors Corporation.

BB, A-1 Ratings: Ratings given to bonds by private agencies that specialize in evaluating credit risks. Companies usually pay these agencies to have their bonds rated. The ratings are then publicized and have an impact on the yield of the rated bonds. Generally, the higher the alphabet, the poorer the credit risk, that is, an A rating is better than a B rating, and an AA rating is better than an A rating.

HLT: Highly Leveraged Transaction (See Chapter 6).

Introduction

Banking used to be a simple business. A bank borrowed money and loaned to others at a spread over cost. The borrowing and lending activities were reflected on the bank's balance sheet.[1] But now banks are as likely to do this business "off-balance sheet" as "on." Chapter 8 discussed some off-balance sheet activities of banks. We continue that discussion with an examination of securitization and loan sales.

When a bank sells a loan commitment, for instance, it needs to provide funding only if the customer exercises the commitment. If a "takedown" occurs, the loan appears on the balance sheet. But the bank can avoid funding, even at this stage, by selling the loan to another bank (a loan sale) or by securitizing it. Securitization involves combining the loan with others of similar characteristics, creating credit-enhanced claims against the cash flows of this portfolio, and then selling these claims to investors.[2]

The practice of loan sales by bank, which we covered in Chapter 7, is quite old; it predates 1880. Securitization, by contrast, is more recent, dating back to 1970 when the Government National Mortgage Association ("Ginnie Mae," or GNMA) developed the GNMA *pass-through*, a mortgage-backed security collateralized by Federal Housing Administration (FHA) and Veterans Administration (VA) single-family mortgage loans. Thus, the S&L industry has been involved in securitization for about 25 years. Banks, on the other hand, are relative newcomers to this market. Although in 1977 Bank of America issued the first private-sector pass-through, which was backed by conventional mortgages, the securitizing of various types of bank loans did not begin until 1985.

This market, often known as the market for Asset-Backed Securities (ABS), had grown to almost $2 trillion by the end of 2005. The origins of the ABS market can be traced to familiar lending practices such as factoring and secured lending, and the market subsequently evolved to the securitization of pools of home mortgages. Nonmortgage asset securitization began in March 1985 when Sperry Lease Financial Corporation floated a $192.5 million public offering. These pass-through securities (which represent direct ownership claims against the securitized portfolio) were secured by a pool of lease receivables originated by Sperry Corporation, now Unisys Corporation. Letters of credit from Union Bank of Switzerland facilitated a triple-A debt rating for the issue.

In the years that followed, securitization increased dramatically (see Figure 9.1 for post-1995 growth). Currently, a wide range of assets are securitized. Examples are: automobile loans and leases, credit card receivables, commercial truck loans, and boat loans. Private issuers include commercial banks, finance subsidiaries of indus-

1. No wonder Walter Bagehot, an economist, said, "The business of banking ought to be simple; if it is hard, it is wrong." [Bagehot (1873)].

2. For a good review of securitization, we recommend Pavel (1986, 1989) and Monahan (1989).

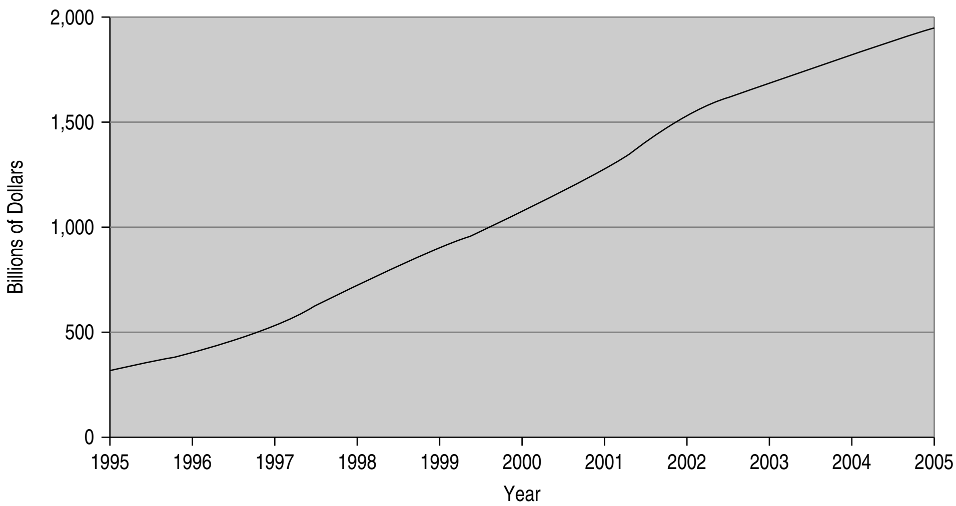


FIGURE 9.1 Total Amount of Asset-Backed Securities Outstanding
Source: The Bond Market Association.

trial companies, and savings institutions. See Table 9.1 for data on different types of securitized assets.

The stated maturities in the ABS market usually do not exceed 6 years and average lives have ranged from 6 months to 5 years. Most of the market is concentrated in the 18-to-36-month period.

An initial obstacle to securitization was uncertainty about whether the Glass-Steagall Act problems on underwriting or distribution of corporate securities also prohibited securitization. However, in the mid-1980s, the Office of the Comptroller of the Currency (OCC) ruled that national banks could sell interests in pools of loans. A court of appeals upheld the OCC's position and ruled against the Securities Industries Association (SIA). The court ruled that sale of asset-backed securities was not limited by Glass-Steagall because these instruments were "not securities but investments in the underlying loans." The Supreme Court later refused to hear an appeal by the SIA, thereby establishing the right of national banks to securitize.[3]

In the rest of this chapter, we cover a fairly wide range of topics pertaining to loan sales and securitization. In the next section we explain securitization and loan sales as natural outcomes of the desire to capture some of the gains from decomposing the traditional lending function. Then we describe the different ways in which securitization is achieved. This is followed by an examination of the economics of securitization in greater detail. Accounting and regulatory issues are examined in the next section. After this we explore the strategic issues faced by banks participating in the ABS market. Loan sales are examined subsequently, and this is followed by the concluding section. A case study is provided to illustrate the strategic securitization issues facing banks.

3. See Huber (1992).

TABLE 9.1 Asset-Backed Securities Outstanding by Major Types of Credit 1995–2005

	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
Total Amount Outstanding	316.3	404.4	535.8	731.5	900.8	1071.8	1281.2	1543.2	1693.7	1827.8	1955.2
Automobile	59.5	71.4	77	86.9	114.1	133.1	187.9	221.7	234.5	232.1	219.7
% of Total	18.8%	17.7%	14.4%	11.9%	12.7%	12.4%	14.7%	14.4%	13.8%	12.7%	11.2%
Credit Card	153.1	180.7	214.5	236.7	257.9	306.3	361.9	397.9	401.9	390.7	356.7
% of Total	48.4%	44.7%	40.0%	32.4%	28.6%	28.6%	28.2%	25.8%	23.7%	21.4%	18.2%
Home Equity	33.1	51.6	90.2	124.2	141.9	151.5	185.1	286.5	346.0	454.0	551.1
% of Total	10.5%	12.8%	16.8%	17.0%	15.8%	14.1%	14.5%	18.6%	20.4%	24.8%	28.2%
Manufactured Housing	11.2	14.6	19.1	25.0	33.8	36.9	42.7	44.5	44.3	42.2	34.5
% of Total	3.5%	3.6%	3.6%	3.4%	3.8%	3.4%	3.3%	2.9%	2.6%	2.3%	1.8%
Student Loan	3.7	10.1	18.3	25	36.4	41.1	60.2	74.4	99.4	115.2	153.2
% of Total	1.2%	2.5%	3.4%	3.4%	4.0%	3.8%	4.7%	4.8%	5.9%	6.3%	7.8%
Equipment Leases	10.6	23.7	35.2	41.4	51.4	58.8	70.2	68.3	70.1	70.7	61.8
% of Total	3.4%	5.9%	6.6%	5.7%	5.7%	5.5%	5.5%	4.4%	4.1%	3.9%	3.2%
CBO/CDO	1.2	1.4	19	47.6	84.6	124.5	167.1	234.5	250.9	264.9	289.5
% of Total	0.4%	0.3%	3.5%	6.5%	9.4%	11.6%	13.0%	15.2%	14.8%	14.5%	14.8%
Other	43.9	50.9	62.5	144.7	180.7	219.6	206.1	215.4	246.8	258.0	288.7
% of Total	13.9%	12.6%	11.7%	19.8%	20.1%	20.5%	16.1%	14.0%	14.6%	14.1%	14.8%

All amounts in billions.
Source: The Bond Market Association.

Preliminary Remarks on the Economic Motivation for Securitization and Loan Sales

Decomposition of the Lending Function

Lending can be decomposed into at least four basic operations: (a) origination (including underwriting), (b) guaranteeing, (c) servicing, and (d) funding. This decomposition was long obscured by the *modus operandi* of financial institutions, which unified these operations. But there is nothing immutable about this unification. For example, suppose a bank were to specialize in the processing of interest rate and credit risk, along with the provision of brokerage services. It could restrict itself to writing letters of credit and loan commitments, avoiding deposits and earning assets altogether.

So why were these lending functions combined in the first place and why are they being decomposed *now*? The reasons are twofold: funding advantages due to the regulatory environment and information technology. Let us consider each in turn.

The Traditional Benefits of Funding Loans

In earlier times, depository institutions enjoyed an advantage in funding, and they consequently developed the expertise needed to originate and underwrite assets including loans. The funding advantage was a consequence of regulation: deposit interest rate ceilings, underpriced governmental deposit insurance, entry restrictions, and various tax advantages—particularly those related to loan-loss reserves, mutuality, and housing. The resulting rents were shared among depositors, borrowers, and owners/managers of banks and thrifts. This system, introduced in the 1930s following more than a decade of socially disruptive bank failures, was based on an *implicit contract* between depositors, owners/managers of banks and thrifts, and the government. Depositors agreed to accept a below-market return for their funds in exchange for a government guarantee; the guarantee (deposit insurance) transformed bank and thrift liabilities into contingent claims against the U.S. government. Bankers agreed to accept regulation and supervision in exchange for a subsidy that lowered the cost and extended the duration of deposits. The government accepted a residual exposure (on behalf of the taxpayers) under the deposit guarantee in exchange for the political gains from stability in the banking system.

The Erosion of Funding Benefits and the Incentives for Securitization and Loan Sales

The implicit contract between depositors, depository institutions, and the government remained intact until the inflation of the 1970s increased the opportunity cost of deposit holding from something on the order of 100 basis points to 400, 500, or even 600 basis points. This caused depositors to turn to higher-yielding money-market funds. The implicit contract began to unravel.

The trend continued with the legislatively mandated dismantling of deposit interest rate ceilings in the 1980s. As deposit interest rates rose, deposit rents of banks and thrifts eroded. In addition, entry barriers into banking began to crumble, tax preferences began to vanish, and the price of deposit insurance increased and capital requirements were raised. In varying degrees, all of these changes diminished the rents available to banks and the advantages that they enjoyed in funding loans with deposits. However, the originating, monitoring, and servicing skills that they had developed earlier remained intact. This provided the first impetus for banks and thrifts to originate and underwrite loans, but not to fund them, that is, to either sell or securitize loans.

A second impetus for loan sales and securitization was provided by advances in information technology. A successful loan sale requires that the buyer (usually another financial institution) be able to assess the payoff attributes of the loan, which in turn is facilitated by good information. This is even more critical for securitization in cases where the buyers are investors as opposed to financial institutions. Improvements in information processing technology have made it easier for investors to rate assets, and therefore reduce informational gaps between investors and the originators of loans (banks). Moreover, information technology has been the key to the servicing and monitoring provided by financial institutions, especially with stripped cash flows. This has facilitated securitization.[4] This argument can be seen quite clearly in the (somewhat oversimplified) numerical example given below.

4. See Greenbaum (1987) and Greenbaum and Thakor (1987) for a discussion of securitization that assigns a role to information processing costs. See also Kareken (1987) and Fishman and Kendall (2000). An examination of the effects of asset securitization appears in Thomas (2001).

Example 9.1 Suppose the North American Bank has originated a portfolio of loans. North American knows that the aggregate payoff on this portfolio will be \$100 with probability 0.9 and \$30 with probability 0.1. Call this portfolio A. Investors, however, are unable to distinguish between this portfolio and another loan portfolio, call it B, that has an aggregate payoff of \$100 with probability 0.7 and \$30 with probability 0.3. Investors believe that there is a 0.5 probability that the portfolio is A, and an equal probability that it is B. There is universal risk neutrality.

The cost to the bank of communicating the "true" value of its loan portfolio is \$11; this can be viewed simply as a charge against the revenue from the sale of the loan portfolio. Think of this as a signaling cost (Chapter 1) that declines with advances in information technology because these advances enable firms to resort to lower-cost signaling mechanisms. Also, North American's net profit from loan origination and servicing is 1 percent of the value of the securitized loan portfolio, whereas if the loans are kept on the books and funded by the bank, the bank's net profit is 2 percent of the "true" value of the loan portfolio minus a fixed cost of 99 cents associated with funding (this could represent, for instance, the sum total of regulatory taxes and administrative costs). Will North American prefer to securitize or fund its loan portfolio? Does your answer change if the communication cost drops from \$11 to \$2?

Solution We will solve this problem in three steps. First, we will show that if North American decides to sell/securitize its loan portfolio, it will prefer to do so *without* communicating information to investors since the cost of communication exceeds the benefits of revelation. Having shown that securitization without communication dominates securitization with communication, in step 2 we show that North American prefers to fund the loan rather than securitize it without communication. Finally, in step 3 we show that North American prefers to securitize with communication if the communication cost drops from \$11 to \$2.

Step 1 First, we compute the value of the "pooled" portfolio, that is, the price at which the bank can sell or securitize the portfolio without any information communication. Given risk neutrality, the bank offering portfolio A will be able to sell it for the average of the values of portfolios A and B, that is, at

$$\underset{\text{(expected value of loan portfolio A)}}{0.5[0.9 \times 100 + 0.1 \times 30]} + \underset{\text{(expected value of loan portfolio B)}}{0.5[0.7 \times 100 + 0.3 \times 30]}$$

$$= 0.5[93] + 0.5[79] = \$86.$$

Then, it is apparent that it does not pay for North American to reveal its true portfolio quality to investors, since its net payoff from doing so is \$93 (the privately known value of its loan portfolio) minus \$11 (the cost of information communication), which equals \$82, whereas the "pooling value" of its loan portfolio is \$86. Thus, securitization without communication dominates securitization with communication.

Step 2 You can see now that if North American securitizes its portfolio without communication, its net profit is 86 cents (1 percent of \$86). But if it funds the loans, its net profit is $0.02 \times \$93 - 0.99 = 87$ cents. Thus, the bank will prefer *not* to securitize when the cost of communicating the true value of its loan portfolio to investors is \$11. Combining steps 1 and 2 shows that funding the loan is North American's optimal strategy.

Step 3 If the communication cost drops to \$2, North American's net profit from communicating and securitizing is $0.01 \times \$[93 - 2] = 91$ cents. This exceeds both the net profit from funding the loans as well as the net profit from securitizing without communication, and shows how improvements in information processing technology that reduce the costs of communicating financial information—can spur securitization.
A more complete discussion of this phenomenon appears in a later section.

Different Types of Securitization Contracts

Loan-backed securities are collateralized by residential, multifamily, and commercial mortgage loans, automobile loans, credit card receivables, Small Business Administration loans, computer and truck leases, loans for mobile homes, and various finance receivables. There are three basic types of asset-backed securities, each of which evolved from the secondary mortgage market.

Pass-Throughs

The first type of loan-backed security is a *pass-through*, which represents *direct ownership* in a portfolio of mortgage loans that share similar maturity, interest rate, and quality characteristics. The portfolio is placed in a grantor trust and certificates of ownership are sold directly to investors; each certificate represents a claim against the entire loan portfolio. The loan originator (say a bank or a thrift) services the portfolio and collects interest and principal on the loans, although sometimes origination and servicing are provided by different institutions. The servicer deducts a fee from the collected proceeds and passes the difference along to the investors; hence the name "pass-through." Ownership of the loans (mortgages) rests with the certificate holders. Thus, pass-throughs do not appear on the originator's balance sheet. There are two structures used with pass-throughs: static pool and dynamic pool. Each is discussed below.

Static Pool Pass-Throughs: The term "static" here refers to the nature of the pool of loans against which claims are sold to investors; this pool is fixed. The trust in which the loans are held is tax free at the trust level. Taxes are levied at the *beneficiary* level. Most pass-through securities provide for monthly payments of principal and interest. Figure 9.2 shows a schematic for a typical static pass-through structure.[5] The payments made by borrowers are paid into a separate interest-bearing account maintained in the trust department of an insured bank (the trustee) in the name of the trustee. This account is known as the *collection account*. Payments into this account are applied first to pay a monthly servicing fee. On each payment date, the trustee passes along the monthly payments of principal and interest to investors. The servicer is responsible for paying the trustee's fee.

There is usually *credit enhancement* of the loan portfolio. This enhancement is provided by posting "excess" collateral and/or through an *insurance bond* purchased

5. The ensuing is based in part on Pavel (1989).

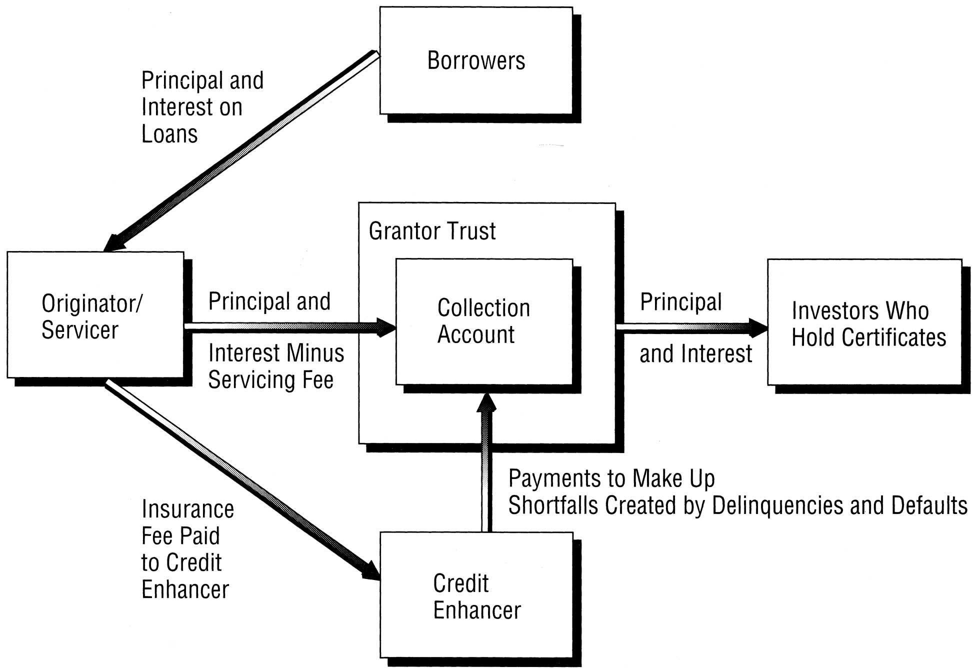


FIGURE 9.2 Cash-Flow Schematic for a Static Pass-Through

by the originator. The protection covers some proportion of the underlying assets at the date of issue. For example, suppose there is 15 percent credit enhancement with an insurance bond. The credit enhancer is then responsible for loan defaults up to that percentage of the value of the securitized loan portfolio.[6] In effect, the credit enhancer purchases the defaulted contracts. With a credit enhancement, the guarantor trust is entitled to payments from the credit enhancer to cover the losses of the loan portfolio due to defaults up to the specified coverage.

The most common type of static pass-through is the Ginnie Mae, which is a mortgage-backed security collateralized by FHA-VA mortgages. The GNMA, a direct agency of the federal government, acts as a credit enhancer, guaranteeing timely payment of principal and interest. Thus, these pass-throughs are virtually free of default risk for investors. A highly developed secondary market ensures liquidity for these instruments. The federal Home Loan Mortgage Corporation (Freddie Mac), an indirect agency of the federal government, developed a similar pass-through security in 1971, called the "participation certificate" (PC). The Federal National Mortgage Association (FNMA, or Fannie Mae) developed the mortgage-backed security (MBS) in 1981. Both the PC and the MBS are backed by portfolios of uninsured and privately insured mortgage loans. Monthly interest and full repayment of principal on PCs are guaranteed by Freddie Mac, but the timing of principal payments is not.

6. The level of credit enhancement is typically determined by the credit-rating agency as the minimum that is required in order to assign to the loan portfolio the credit desired by the issuer.

Private sector pass-throughs are less common than these federal agency issues. In 1977, Bank of America issued the first private sector pass-through. These securities were backed by conventional mortgages, and private mortgage insurance was purchased to cover the entire pool of loans rather than each individual loan. Since the insurance covered the loan portfolio as a whole, diversification available to the insurance company meant lower insurance cost than if individual loans, representing a subset of the portfolio, had been insured.

Dynamic Pool Pass-Through Structure: "Dynamic" refers to the pool of loans against which claims are sold to investors. The debt obligations included in the pool are usually *short term*, so that they turn over, implying changes in the composition of the loan portfolio. This structure, also known as a "revolving structure," involves a pool of loans with an average life that is *shorter* than the stated maturities of claims issued against the pool. When a loan within the pool matures, the proceeds are reinvested for a fixed period of time (the "revolving period"). During the revolving period (the duration of which can be structured to satisfy desired asset considerations), only interest is paid to the certificate holders. All principal repayments are reinvested to maintain the original principal amount. Principal amortization begins at the end of the revolving period, usually on a pass-through basis. This design is most often used with credit card receivables (for example, JCP Master Credit Card Trust for JC Penney credit card receivables and Sears Credit Account Trust for Sears credit card receivables) where repayment periods are uncertain and can be very brief, frustrating the desire of investors to remain invested for some minimum period.

Asset-Backed Bonds

The second type of ABS is the asset-backed bond (ABB). Like the static pass-through, the ABB is collateralized by a portfolio of loans. The main difference is that in the case of an ABB, the originator sells the assets to a *wholly owned subsidiary* created for the sole purpose of securitizing the assets. Consequently, the *assets remain on the originator's (consolidated) balance sheet*. That is, instead of selling the assets to a trust that subsequently sells claims against the assets to investors, the subsidiary itself issues claims (general obligation notes) to investors. These claims are secured solely by the assets of the subsidiary and any credit enhancement obtained for the purpose. Figure 9.3 depicts the cash-flow structure of an ABB. As the figure indicates, the finance company, which is a wholly owned subsidiary of the originator, issues certificates/notes to investors, usually through an investment bank that underwrites the issue. The revenues collected by the finance company from principal and interest payments are transferred to a trustee. These revenues are added to cash contributions made by a credit enhancer and then disbursements are made to investors by the trustee.

An important difference between a pass-through and an ABB is that the cash flows from the pool of assets that serve as collateral are *not* dedicated to the payments of principal and interest on ABBs. The maturity on an ABB is usually prespecified (normally 5 to 12 years) and interest is generally paid semiannually.

ABBs are usually *overcollateralized*. The norm is to evaluate the collateral quarterly, and to augment it if the value falls below an amount stated in the bond indenture. The reasons for overcollateralizing are twofold.

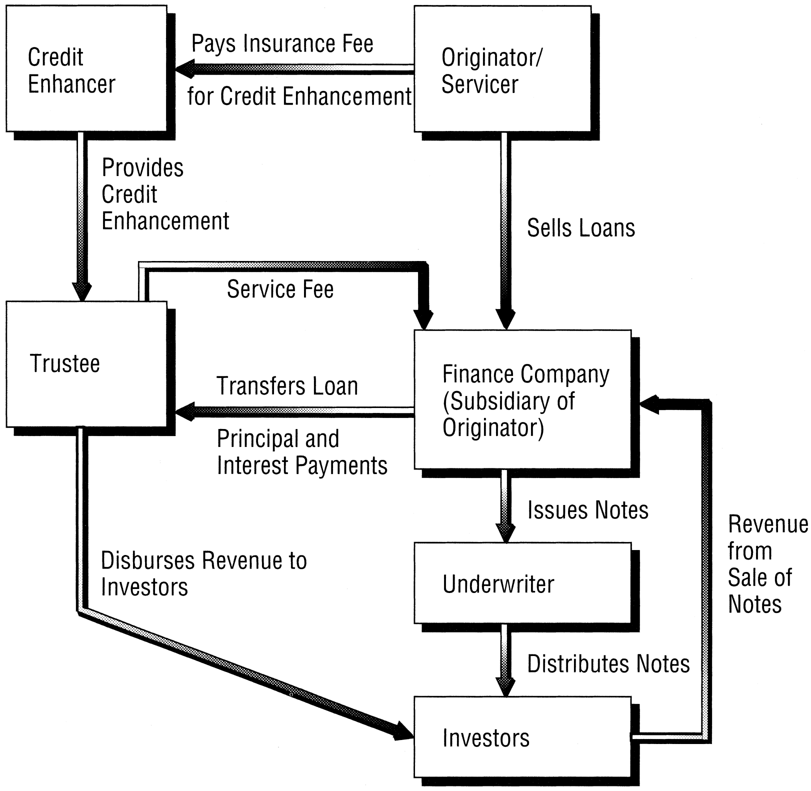


FIGURE 9.3 Cash-Flow Schematic for an Asset-Backed Bond

1. Overcollateralization, like other forms of credit enhancement, increases safety for investors and therefore reduces the required yield on the ABB. This benefits the originator because any proceeds beyond what is needed to service the principal and interest on the ABB accrue to the originator but are used to augment the value of the collateral pool. Thus, the overcollateralization is a particular form of *reinvestment* of the proceeds collected by the originator from the pool of assets. By reinvesting to increase the collateral, the originator reduces the risk faced by investors in the same way that a borrower reduces a lender's risk by using project cash flows to purchase additional collateral rather than using the cash flows to increase dividends to shareholders. That is, overcollateralization improves the overall outcome by diminishing a form of *moral hazard*. Other forms of credit enhancement, such as an insurance bond, also diminish moral hazard, but work a little differently. For example, with an insurance bond, the firm providing the bond would be expected to monitor the originator to ensure sufficiently high asset quality.
2. Excess collateral also protects investors against decreases in the market value of the collateral between valuation dates. This is a simple *risk-sharing* argument. If the originator is more risk tolerant than individual investors—perhaps because of a superior ability to hedge risks—then the originator provides investors a form of value-fluctuation insurance in exchange for a lower yield on the ABB. In this sense, overcollateralization is no different from any other form of credit enhancement.

ABBs have been used by both public and private entities, but private issues dominate. Major private issuers are savings and loan associations and mutual savings banks. The ABB market, however, is much smaller than the market for pass-throughs (about 5 percent of the market for pass-throughs). One reason for this may be that an ABB stays on the originator's books. Thus, a financial institution must hold both reserves and capital against an ABB.

Pay-Throughs

The third type of ABS is the *pay-through bond*. This security combines features of the pass-through and the ABB. Its similarity to the ABB is that the pay-through appears on the originator's balance sheet as debt. Its similarity to the pass-through is that cash flows from the pool of assets used as collateral are dedicated to servicing the bonds.

CMO: In June 1983, Freddie Mac issued a pay-through bond called a *Collateralized Mortgage Obligation* (CMO). Each CMO issue was divided into three "tranches" (maturity classes), and each class received semiannual interest payments. The tranches, however, were strictly prioritized for the receipt of scheduled principal payments and repayments. That is, Class A bondholders received the first installments of principal payments, and any prepayments, until Class A bonds were paid off. After Class A bondholders were paid off, Class B bondholders began to receive principal payments and repayments. Class B bondholders have to be completely paid off before Class C bondholders could. Class A bondholders were repaid within 5 years of the offering date, Class B bonds will be repaid within 12 years, and Class C bonds within 20 years. Figure 9.4 provides a cash-flow diagram for a CMO for the first five years or until tranche A is repaid.

Under the structure in Figure 9.4, Class A would receive interest plus all of the principal payments passed through from the underlying mortgages until it is entirely paid off (which is estimated to take no more than five years), while Classes B and C receive only interest. While there is still variability in the rate of repayment due to the randomness in prepayment rates, the CMO structure reduces this variability by "serializing" cash flows this way. That is, CMO-holders receive a kind of "call protection." They can be reasonably confident that their bonds will not prepay (be called) prematurely.

CMOs facilitate the management of *prepayment risk*. Routinely borne by financial institutions, this risk arises from the fact that borrowers tend to *prepay* their debts when interest rates fall because they can refinance at lower rates. This is especially true for long-term mortgages that have no prepayment penalties. Thus, the financial institution does not fully benefit from the decline in its cost of funds relative to the rates on its longer-maturity assets. On the other hand, when interest rates rise, the institution's cost of funds rises but its asset returns do not, as borrowers hold on to their low-interest-rate mortgages. By investing in a CMO tranche with a sufficiently long effective maturity (say Class C in Figure 9.4), the institution can reduce its exposure to prepayment risk.

A wide variety of CMOs are available. They range from three maturity classes to more than six. Most, however, have four tranches, which include three "regular" maturity classes and a "residual" class, called the "Z class." The first three classes

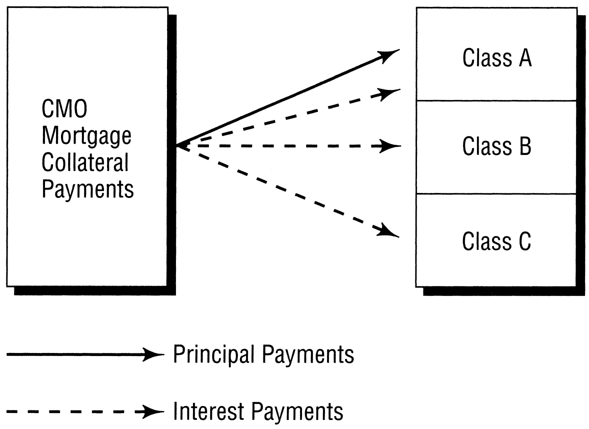


FIGURE 9.4 Cash-Flow Schematic for a CMO (for the First Five Years)

are paid interest at the stated rates, starting with the issue date. The Z class is basically an accrual bond in which earned interest accrues to the principal and is compound while the other classes earn interest. After the first three classes are paid off, the Z class receives regular principal and interest payments along with accrued interest.

CMOs can be used by financial institutions to facilitate asset/liability management.[7] For example, suppose an S&L has 30-year fixed-rate mortgages financed with liabilities of shorter maturities. Such an institution could reduce the maturity mismatching on its balance sheet by swapping its mortgages for the shorter tranches of CMOs.

Both public and private sector firms participate in the CMO market. Issuers include investment banks, federal agencies, builders, and thrift institutions, among others. The issuance of a CMO is *not* an asset sale because the debt obligation stays on the originator's books. This structure was adopted to comply with tax regulations that stated that a trust could not qualify for *grantor trust status* if it issued multiclass claims that divide the cash flows in a *not prorata* fashion, as a CMO does; recall that this "tax problem" does not arise with pass-throughs because certificate holders do receive the cash flows in a prorata fashion. Thus, a CMO had to tolerate the "inefficiency" of keeping the collateral assets on the originator's balance sheet rather than selling these to a trust. This resulted in another tax disadvantage, namely that due to regulatory taxes like reserve and capital requirements and deposit insurance premiums. This limited the use of CMOs.

REMICs: The Tax Reform Act (TRA) of 1986 authorized REMICs (Real Estate Mortgage Investment Conduits). The main difference between a CMO and a REMIC is in the *tax treatment*. REMICs can qualify as *asset sales* for *tax purposes* if the following conditions are satisfied.

7. See, for example, Kasper (1985).

1. A REMIC must contain at least one regular class and no more than one residual class.
2. The collateral of a REMIC must consist of "qualified mortgages" or "permitted investments." Qualified mortgages include single and multifamily residential mortgage loans and commercial mortgages as well as mortgage-backed securities. Permitted investments include short-term interest-bearing securities used only for reinvesting monthly cash flows prior to their scheduled transfer to bondholders, investments to fund operating expenses of the REMIC, and properties acquired through foreclosure.

Securitization Innovations

New types of securitization contracts continue to proliferate for three main reasons. First, lower interest rates have made the prepayment options in mortgages more valuable to investors, who have become more sophisticated in dealing with prepayment risk. Thus, new securities that facilitate the management of prepayment risk have been created.[8]

Second, investors who are relatively uninformed about the probability distributions of future payoffs on various securities will find themselves at a disadvantage in dealing with investors who are better informed. The uninformed will therefore demand relatively information-insensitive securities that would enable them to trade without being expropriated. The cash-flow stripping that accompanies securitization often creates information-insensitive securities by partitioning the cash flows of a composite, information-sensitive security in such a way that the senior-most security is a nearly riskless bond that would appeal to uninformed investors.[9]

Third, innovations in securitization and cash-flow stripping have also facilitated the creation of securities that appeal to informed investors. A given security with some private-information content can always be stripped into two securities, one of which is more information sensitive (has a greater private-information content) than the original security. Those who have the ability to acquire information at a cost will find that the return on their investment in information is greater with the more information-sensitive security. Thus, more of these investors will become informed, and there will be a greater demand from these investors, which in turn will elevate the price at which the security trades, and hence the issue's revenue.[10] A similar argument can be made for other attributes of the original security, which can be altered through cash-flow stripping to appeal to specific clienteles of investors. For example, investors may have different cash-flow preferences due to tax considerations and risk attitudes. Then, stripping a security into components permits the issuer to cater more effectively to the desires of these clienteles than issuing only a single class of pass-throughs would permit. Translated into dollars, this means higher revenue for the issuer. Investors desire a low prepayment rate even when interest rates are falling. Investors who want a discount mortgage security with higher prepayment can purchase a P/O ("principal-only" security).

8. See Dougherty (1987).
9. Gorton and Pennacchi (1990) have proposed this explanation as a way to understand the preponderance of diversified baskets of securities, riskless CDs, and other information-insensitive assets.
10. This intuition is provided by Boot and Thakor (1993).

Stripped Securities: The *stripped mortgage security* (the "strip") involves two classes of pass-through securities that receive different portions of principal and interest from the same pool of mortgage loans. For example, a pool of mortgage loans with an average APR of 8 percent might be split into a "premium" security with a 12 percent coupon and a "discount" security with a 4 percent coupon. When this process of "stripping" is taken to its logical extreme, it creates interest-only (I/O) and principal-only (P/O) securities. This is known as an I/O–P/O *strip*. Holders of the I/O strip receive primarily interest payments from the securitized pool, whereas the holders of the P/O strip receive nearly all of the principal payments.

Strips offer advantages to issuers as well as to investors. The advantages to investors are based on the clientele's argument presented earlier. That is, some investors may prefer information-insensitive securities, others may prefer information-sensitive securities, and yet others may desire a specific cash-flow pattern due to tax considerations. Strips can satisfy these different demands.

A financial institution can also use I/O–P/O strips for hedging against interest-rate risk. I/Os are useful in hedging fixed-rate mortgage loans and other fixed-income assets. An increase in interest rates causes a decline in the value of the I/O as with other fixed-income assets. However, the higher interest rates will slow down prepayments. This will generate higher-than-expected cash flows for the holders of the I/O strips, and will thereby increase the value of the I/O. In most interest-rate scenarios, an increase in interest rates will cause prepayments to fall sufficiently to create an *inverse relationship* between the value of the I/O strips and the prices of the bank's other fixed-income assets. This provides hedging.

A financial institution can use P/O strips to hedge its fixed-income liabilities. A decrease in interest rates increases the value of a P/O strip because the discount rate for computing the present value of future principal payments has decreased. Moreover, prepayments go up, accelerating the cash flows accruing to the holders of the P/O strips, further increasing the value of that strip. Thus, the value of a P/O is inversely related to the values of fixed-income liabilities, and hedging is possible.

Asset-Backed Commercial Paper (ABCP): ABCP is commercial paper secured by designated corporate assets, typically receivables. The term ABCP is almost oxymoronic since commercial paper is understood to be the traded, short-term *unsecured* debt of corporations. Figure 9.5 illustrates how an ABCP program works.[11]

A bank establishes a "special purpose corporation" (SPC). The SPC purchases credit card receivables or other assets from a corporation (the seller) in need of funding. To finance this purchase, the SPC issues commercial paper that is secured by the assets purchased by the SPC. The bank provides credit enhancement enabling the SPC to obtain a high credit rating for the commercial paper, typically through overcollateralization and/or a standby letter of credit. This credit enhancement enables the SPC to obtain a high credit rating for the commercial paper.

The ABCP market, which came into existence in 1983, has expanded rapidly, growing to a multibillion-dollar market within a decade. Fitch Investors Service reports that, in mid-1993, 175 ABCP programs were in operation, representing \$75 billion in outstanding commercial paper and \$150 billion in commitments. At May 2006, there was over \$900 billion in outstanding ABCP. Global banks like ABN Amro Bank (the Netherlands), Societe Generale (France), Deutsche Bank (Germany), Barclays Bank (United Kingdom), National Westminster Bank PLC

11. This discussion is based in part on Cutler and Sveen (1993) and Kraus (1993).

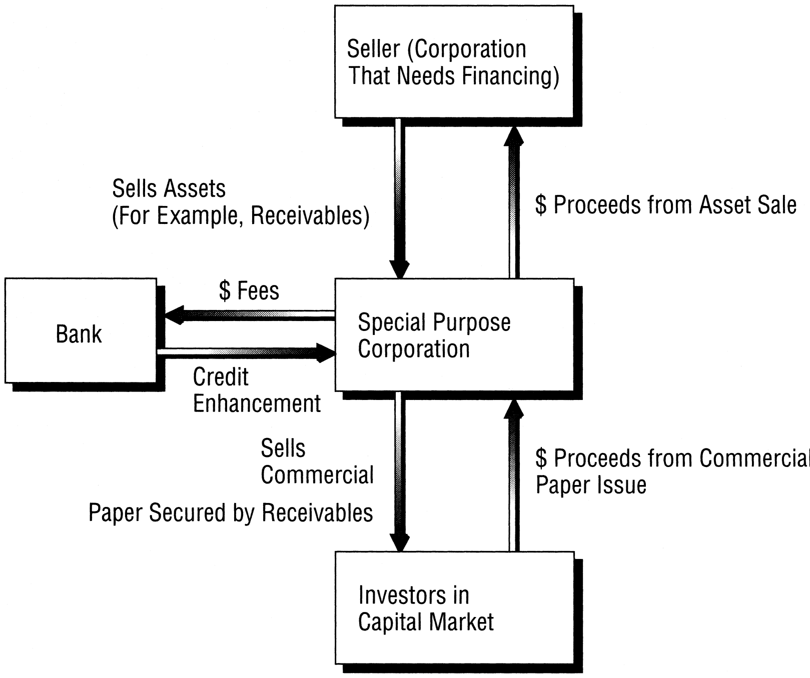


FIGURE 9.5 A Typical ABCP Program

(United Kingdom), Industrial Bank of Japan, Sumitomo Bank (Japan), and Imperial Bank of Commerce (Canada) represent a group whose presence in this market is growing rapidly. The typical size of an ABCP issue by a foreign bank is $1 billion.

Why has ABCP grown so in popularity? We will examine both the demand and supply sides. On the demand side, ABCP offers some firms lower cost funding than either "regular" (unsecured) commercial paper or a bank loan. Regular commercial paper may either be unavailable or too costly because of the high cost of moral hazard owing to the unsecured nature of the paper (recall Chapters 5 and 6). A bank loan may be too costly because of capital and reserve requirements. One reason why ABCP lowers the firm's funding cost is the credit enhancement provided by the bank. Not only does this directly lower the investor's risk in holding the paper, but it also signals the bank's involvement in monitoring the borrower[12] and is certification that the borrower is creditworthy. Thus, the basic screening and monitoring services provided by the bank play a key role in the ABCP market.

On the supply side, the risk-based BIS (Bank for International Settlements) capital rules appear to have increased the benefits of ABCP to banks. If the bank were to extend a loan to the borrower, it would not only need to keep reserves against the deposit used to fund the loan, it would also need to set aside capital equal to 8 percent of the loan amount. With an ABCP, the bank would need to keep capital equal to 8 percent of only the credit enhancement (typically a fraction of the total borrowing). For example, a $1 billion bank loan would need $80 million in bank

12. The credit enhancement strengthens the bank's incentive to monitor the borrower because the bank has more to lose if the borrower defaults.

capital, but with an ABCP program, a bank might issue a letter of credit equal to only 10 percent of the total amount, so that only $8 million in capital would have to be set aside.[13] Thus, as with other off-balance sheet products, banks are able to earn fee income without posting as much capital as with conventional funding.[14]

Securitization of Other Assets: CARS, CARDS, Intellectual Property and So On

Auto Loans: Securitization of automobile loans began in 1985,[15] and from 1985 to 1987, it was the largest sector of the ABS market.[16] By 2005, auto-loan securitization had reached almost $220 billion. The securitization of auto loans is actually the securitization of retail installment sales contracts that are backed by autos and light trucks. The maximum maturity of the loan is 60 months and the loans pay principal and interest on a monthly basis. These loans are packaged and sold as loan-backed securities called *Certificates of Automobile Receivables*, or "CARS," in keeping with Wall Street's penchant for catchy acronyms. They are usually *pass-through* securities, with both the principal and interest passed on directly to the certificate holders. However, the pay-through structure has also been used (for example, by GMAC).

CARS usually involve a higher servicing fee than mortgage-backed securities because an auto loan requires more monitoring. Moreover, the value of the collateral (the car) tends to depreciate somewhat unpredictably through time, compared to the value of a home. Nevertheless, auto loans are readily securitizable because they have predictable default rates, as well as reasonably stable prepayment rates.

In the securitized auto-loan market prepayment speed is usually indicated by the "absolute prepayment rate." This rate represents the percentage of the original loans that are expected to prepay every month. For example, a 2 percent rate means that 2 percent of the original number of the loans in the pool can be expected to prepay every month. The prepayment speed is estimated prior to the offering of the CAR and is a key factor in its pricing. Figure 9.6 gives an example of the cash-flow characteristics of a CAR over its life.

Credit Cards: Securitization of credit-card receivables began in April 1986 when Salomon Brothers privately placed $50 million of pass-through backed by a pool of Bank One credit-card receivables. These securities were called *Certificates of*

13. Noting the disadvantage of funding, Mr. Joseph Rizzi, vice president for structured finance at ABN Amro in Chicago in 1993, said, "Pricing for credit has deteriorated, and the Bank for International Settlements capital rules make no distinction between lending to a triple-A company and a hot dog stand." [Kraus (1993)].

14. Kraus (1993) quotes Mr. James Carson, managing director for asset securitization at Canadian Imperial Bank of Commerce, "Our earnings from this business (ABCP) have grown from zero to a fairly nice number."

15. The first public offering of securitized auto loans was in March 1985. Salomon Brothers offered $60 million of pass-throughs backed by auto loans that were originated and serviced by Marine Midland Bank. A private insurer insured the pool, and a trust was established to hold the underlying loans.

16. See Monahan (1989).

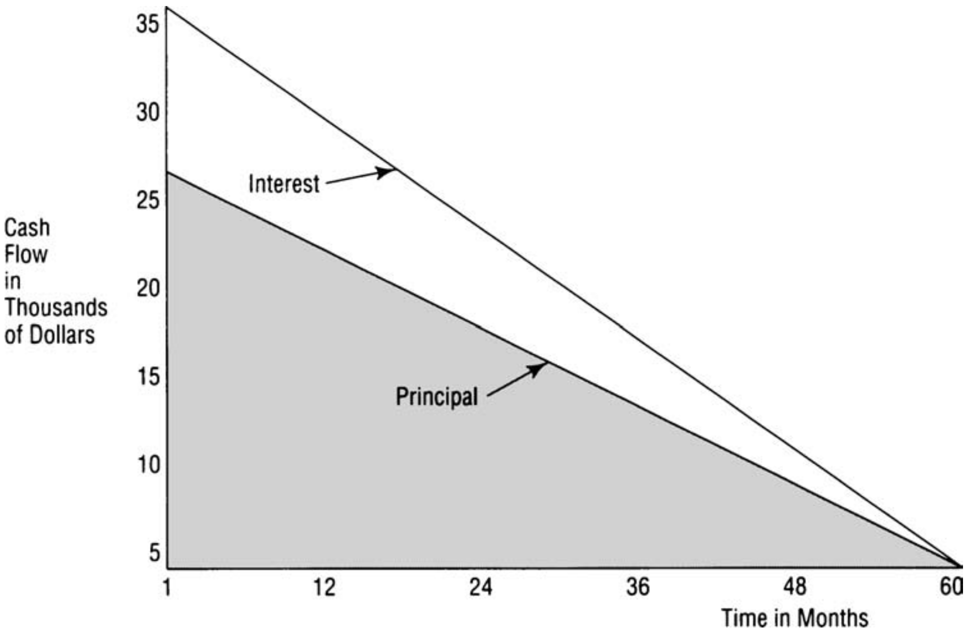


FIGURE 9.6 Cash-Flow Breakdown for CAR (Assuming an Absolute Prepayment Rate of 1.3 Percent)
Source: Monahan, Maureen, *An Investor's Guide to Asset-Backed Securities*, Shearson Lehman Hutton, Inc., March 1989.

Amortizing Revolving Debts or "CARDS." The original CARDS had a stated maturity of 5 years. For the first 18 months, only interest payments were passed through to investors. Principal payments received during this time by the grantor trust were used to purchase additional receivables. Investors began to receive principal payments after the first 18 months. These CARDS were not guaranteed by a third party. Bank One provided credit enhancement through excess collateral by establishing a reserve fund equal to twice the historical default rate on credit-card debt. The bank also retained a 30 percent interest in the credit-card pool and recouped whatever was left in the reserve fund after covering defaults. This "reserve fund" concept has been applied to securitizations of assets other than credit cards, such as CARS.

Following the Bank One experiment, other banks entered the market. For instance, the Republic Bank of Delaware offered a dynamic version of CARDS with a revolving/pass-through security in January 1987. CARDS overtook CARS in 1988 as issuers like Citicorp and Sears entered the market for the first time. By 2005, the volume of credit-card securitizations had reached about $357 billion.

Outstanding balances on credit cards tend to pay down quickly. This is why CARDS often have a "lock out" or "call-protection" period (such as the 18-month period for the Bank One CARDS) during which only interest is paid to investors. The principal is reinvested in new receivables. Amortization of principal and interest begins after the lockout period. The length of the amortization period depends on the characteristics of the credit-card pool, but this is not difficult to estimate because monthly repayments rates are predictable. Figure 9.7 illustrates the cash-flow structure and suggests why an originator/servicer might be interested in CARDS.

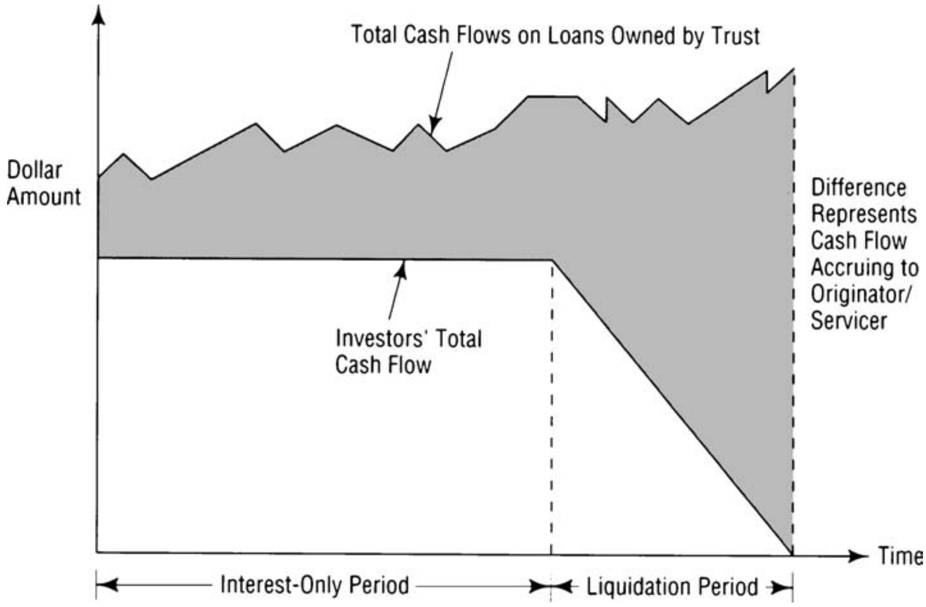


FIGURE 9.7 Pool Cash Flow for CARDS
Source: Monahan, Maureen, *An Investor's Guide to Asset-Backed Securities*, Shearson Lehman Hutton, Inc., March 1989.

The credit-card pass-through structure has undergone many refinements. For example, a *bullet principal payment* structure has been adopted whereby investors receive their entire principal in one lump sum at maturity. Prior to that they receive periodic interest payments exclusively. This provides investors against prepayment risk and early amortization.[17]

Other Tangible Assets: Loans guaranteed by the Small Business Administration (SBA), computer leases, and various types of trade credit have also been securitized. Securities backed by lease receivables[18] and trade credit are similar to mortgage-backed bonds. Commercial paper or corporate bonds are collateralized by lease or trade credit receivables, and the receivables remain on the issuer's balance sheet. Sometimes the issuer sells the receivables to a subsidiary that issues debt collateralized by these receivables.

Finance leases have also been securitized. A finance lease operates over most of the leased asset's useful life. Finance leases either cannot be canceled, or if they are

17. An example of the bullet structure is Citibank's National Credit Card Trust 1989-1. The certificates were backed by VISA and MasterCard receivables and rated AAA/Aaa. The lockout period was 24 months, and in the 12-month period following this, prepayments were reinvested in order to meet principal and interest requirements. Union Bank of Switzerland provided both a maturity guarantee for up to 46 percent of the principal amount and a 12 percent L/C to cover potential credit risk.

18. The securitization of computer leases was first done by Comdisco in 1985, when it sold $35 million in $4\frac{1}{2}$ -year bonds backed by computer leases. In March 1985, Sperry Corporation followed with a $192.5 million issue of 6-year notes backed by computer leases, and in September 1985, Sperry issued another $145.8 million in debt collateralized by computer leases.

cancelable they require the lessee to reimburse the lessor for any losses that may be incurred as a result of the cancellation.

Still other assets that have been securitized include junk bonds, leveraged buyouts, loans for manufactured homes, and commercial loans. Other than trade receivables, securitized commercial loans include loans for employee stock option plans (ESOPS), and leveraged buyout loans.[19]

Securitization is now viewed as a vehicle that can be used for almost any asset. Among "nontraditional" assets that have been securitized are unsold airline seats, song royalties, proceeds from tobacco litigation (being securitized by various states), and natural resources like unsold oil and natural gas. The accounting statute governing securitization in the U.S. is FAS 140.

Securitization of Intangibles Like Intellectual Property

In all of the previous cases of securitization that we have discussed, what was being securitized was essentially a tangible asset. However, conceptually there is no reason why securitization cannot be extended to intangible assets like intellectual property. After all, an important goal of securitization is to enable the borrower to raise cash *now* against the present value of future cash flows, rather than waiting for the future cash flows to materialize. Thus, as long as an intangible asset has the potential to generate future cash flows, one ought to be able to securitize it. And that is precisely what has happened in recent years. The intangible assets that have been securitized include: trademarks, brand names, product designs, corporate name and logo, manufacturing technology, databases and patents.[20] In most cases, the firms involved were seeking financing from additional sources, since they had exhausted debt financing from traditional sources, lacked sufficient inventories to procure additional inventory-backed financing, and real estate was encumbered.

Going Beyond Preliminary Remarks on Economic Motivation: The "Why," "What," and "How Much Is Enough" of Securitization

Why?

Our purpose in this section is to examine in more detail the economics of securitization.

The Supply Side of Securitization: Issuer's Prospective Costs The primary costs of securitization to the issuer are administrative in nature. They include legal fees,

19. Pilgrim Group, Inc., a mutual fund company based in Los Angeles, securitized commercial loans in 1988 when it began to sell shares in a fund that invests in collateralized bank loans made by a money center and large regional banks to domestic companies.

20. See Anson (2005), and Martin and Drews (2005). This is clearly not an exhaustive list.

investment banking fees, and rating agencies' fees. Other costs include the costs of communicating information to investors and the cost of credit enhancement.

Issuer's Prospective Benefits The numerous potential benefits to the issuer include management of interest rate, increased liquidity, and diversification of funding sources. In addition, securitization enables the originator to focus on the origination, servicing, and monitoring of loans and to avoid certain taxes and regulatory costs.

(1) Management of Interest-Rate Risk: By securitizing some of its assets, a bank or a thrift may be able to better manage the interest-rate risk. For example, consider an S&L that holds residential mortgages with an average stated maturity of 27.5 years and mostly fixed rate. On the other hand, 65 percent of the typical S&L's liabilities are time and savings deposits that mature in less than a year. This enormous maturity mismatch creates substantial interest-rate risk for the S&L. To reduce this exposure, the S&L could securitize a pool of mortgages using the pass-through method. This would take these assets off its books and shorten the average maturity of its assets, while still allowing the S&L to service the loans and earn the servicing fee.

Another form of interest-rate risk is prepayment risk, which can also be reduced by securitizing. It can replace the mortgages in its portfolio with pass-throughs and CMOs. This helps the S&L to diversify. Moreover, the CMO helps protect against prepayment risk.

The S&L also could attempt to lengthen the average maturity of its liabilities, which would further reduce the gap in maturities of its assets and liabilities. It can do this by issuing mortgage-backed bonds and pay-throughs. The mortgage loans remain on the S&L's balance sheet but the average maturity of the S&L's liabilities is effectively increased since a mortgage-backed bond has an average maturity of about 5 to 12 years. Of course, the S&L can hedge its exposure through other, more conventional, means such as swaps, options, and futures. But in terms of both transactions costs and overall effectiveness, securitization *may* be a superior alternative in many instances.

(2) Increased Liquidity: Securitization can improve the issuer's liquidity. The obvious reason is that assets that were untraded prior to securitization are traded in active secondary markets after securitization; thus, the issuer holds more liquid assets even if it retains any portion of the securitized portfolio. While the observation itself is correct, it is not terribly insightful. As our discussion of liquidity in Chapter 4 indicated, an asset is liquid if it can be sold quickly without much of a loss relative to its "true" value. This, in turn, rules out a large gap in the information the seller has about the asset's future prospects and the information a potential buyer has about those prospects. Now, active trading improves an asset's liquidity because it provides profit incentives for potential buyers to produce information about the asset. The information is then partly transmitted to others (who may be uninformed) through trading volume, prices, and related parameters.[21] Thus, trading increases the availability of information about the asset in the public domain and hence improves liquidity.

However, before an asset can be introduced for trading at a price that does not impose large *initial* losses on the seller, it must have some measure of liquidity, that is,

21. See Grossman (1981).

the informational gap between buyer and seller should not be too large, Securitization achieves this initial measure of liquidity in two important ways. First, third-party credit enhancement reduces the effect of informational asymmetries between the issuer and investors. Credit enhancement works like a (partial) standby letter of credit in that it *substitutes* a portion of the credit risk of the asset pool with the credit risk of the credit enhancer. If investors have a better knowledge of the credit enhancer (who is likely to have an established reputation) than of the securitized pool of assets, then the relevant informational asymmetry is reduced. Sometimes there is an outright letter of credit provided by a credit enhancer, in which case the informational asymmetry is reduced further.

A second way in which securitization improves liquidity *even prior to trading is* through the pooling of a large number of assets and the subsequent partitioning of portfolio cash flows. To see this, think of a CMO with various tranches, each having its own priority status. The importance of the issuer's private information about the future portfolio returns diminishes as one moves further up the priority ladder.[22] That is, private information is less important for the first tranche (that is, the bond with the first claim against portfolio cash flows) than it is for the portfolio as a whole. Thus, by splitting up the portfolio into tranches, the issuer essentially *distributes* the total private information related to the portfolio across the different tranches in a particular way. This in itself does *not* necessarily reduce the total information asymmetry. But it does make the high priority tranches more liquid (less informationally sensitive) than the portfolio as a whole. The issuer can choose to retain its claim against the most information-sensitive (and least liquid), lowest-priority portion of the portfolio cash flow, and sell off the rest. The issuer needs to fund only the retained portion of the portfolio. The rest is liquefied and securitized. The following illustration clarifies this point.

Example 9.2 Suppose the North American Bank has two loans, each of which is due to be repaid one period hence. The cash flows are independent and identically distributed random variables. Each loan will repay \$100 to the bank with probability 0.9 and \$50 with probability 0.1. However, while North American knows this, prospective investors cannot distinguish this bank's loan portfolio from that of the Southside City Bank, which has the same number of loans, but each of its loans will repay \$100 with probability 0.6 and \$50 with probability 0.4. The prior belief of investors is that there is a 0.5 probability that North American has the higher-valued portfolio and a 0.5 probability that it has the lower-valued portfolio. Suppose that North American wishes to securitize these loans, and knows that if it does so without

(*Continued*)

22. This point has recently been noted by Boot and Thakor (1993). Gorton and Pennacchi (1991) and Subrahmanyam (1991) make a somewhat different observation that is also germane. They note that there is a sort of "information diversification" at work when one assembles portfolios of securities, that is, there is less of a private information problem about an entire portfolio than there is with respect to individual securities in the portfolio. This suggests one more way in which securitization improves liquidity.

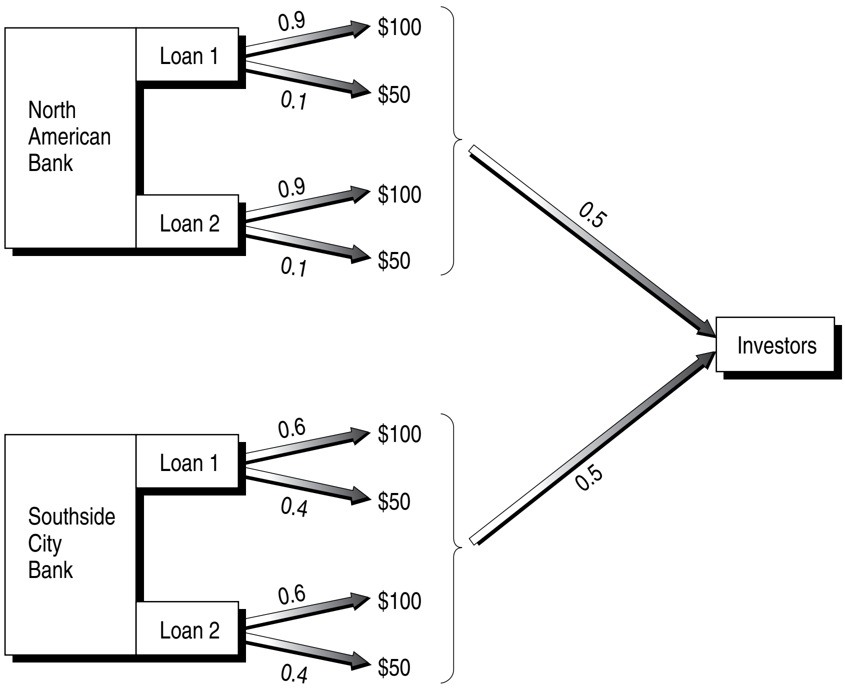


FIGURE 9.8 North American Bank's Loan Portfolio

credit enhancement, the cost of communicating the true value of its loans to investors is 5 percent of the true value. The data for this problem are depicted in Figure 9.8. Explore North American's securitization alternatives. Assuming that a credit enhancer is available and that the credit enhancer could (at negligible cost) determine the true value of North American's loan portfolio, what sort of credit enhancement should North American purchase? Assume investors are risk neutral and that the discount rate is zero.

Solution We solve this problem in four steps. First, we show that if the North American Bank securitizes its loan portfolio as a single security (that is, without cash-flow stripping or creating tranches), it will prefer securitization with communication (investors learn the true value of its loan portfolio) to securitization without communication (in which case investors set a pooling price of the loan portfolio). Second, we examine the benefits of securitizing by creating two tranches represented by two particular classes of bondholders. We show that this cash-flow partitioning benefits North American by increasing its expected revenues relative to the securitization considered in step 1. Third, we stipulate a specific form of credit enhancement. Finally, in step 4 we examine the net benefit of credit enhancement to North American and find that it is positive. Indeed, the loan portfolio is made perfectly liquid by credit enhancement. While credit enhancement by a party that knows North American's portfolio better than investors will always help, it helps to the maximum extent possible here due to the assumption that the credit enhancer can discover the *true* value of North American's loan portfolio at negligible cost.

Step 1 If North American does not communicate any information to investors, the market value of its securitized loan portfolio will be the average value assessed by investors:

{2 total number of loans	$\times$	0.5 probability that loan portfolio is high-valued	$\times$	$[0.9 \times 100 + 0.1 \times 50]\}$ expected value of high-valued loan portfolio
+{2 total number of loans	$\times$	{0.5 probability that loan portfolio is low-valued	$\times$	$[0.6 \times 100 + 0.4 \times 50]\}$ expected value of low-valued loan portfolio

$= \$175.$

Now, North American knows privately that its loan portfolio is worth

$$2[0.9 \times 100 + 0.1 \times 50] = \$190.$$

Thus, if it wishes to communicate its private information to investors, it will cost North American $0.05 \times 190 = \$9.5$. It will then be able to sell its portfolio for \$190 and its net payoff will be $\$190 - \$9.5 = \$180.5$. This means that North American will prefer securitization with communication to securitization without.

Step 2 Consider next the following securitization alternative. North American can create two classes of bondholders in a "senior-subordinated structure" or a "junior-senior structure." Class A bondholders, who receive the first tranche, are entitled to \$100 in the aggregate. After they are paid off, Class B bondholders are entitled to receive \$100 or the residual cash flow, whichever is smaller. Now, since the Class A bondholders will receive \$100 for sure, regardless of whether the loan portfolio is high-valued or low-valued (note that the lowest payoff on either portfolio is \$50), there is no need for the bank to communicate information to these bondholders. The price at which this portfolio can be sold is \$100. Since the Class B bondholders are entitled to receive a maximum of \$100, and the maximum total payoff on the loan portfolio is \$200, it is apparent that these bondholders are essentially residual claimants who receive all of the cash flow remaining after the Class A bondholders are paid off. Thus, the true value of the Class B bonds must be equal to the total value of the loan portfolio minus the aggregate value of the Class A bonds, or $\$190 - \$100 = \$90$. Since the market value of the total loan portfolio is \$175 and the market value of the Class A bonds is \$100, the market value of the Class B bonds should be $\$175 - \$100 = \$75$. If North American now chooses to communicate the true value of Class B bonds to investors, it will be able to sell these bonds for \$90, but the communication cost to the bank will be $0.05 \times 90 = \$4.5$. Thus, its net payoff on securitizing this way will be

$$\$100 + \$90 - \$4.5 = \$185.5.$$

(*Continued*)

North American's net revenue is higher when it uses securitization to partition the total cash flow from the loan portfolio into two classes with different "information sensitivities."

Step 3 Now consider credit enhancement. The best way to structure the credit enhancement is to ask the credit enhancer to pay the Class B bondholders the difference between the promised amount of $100 and the actual residual cash flow after the Class A tranche is paid off. Ignoring the possibility of default by the credit enhancer, this guarantees that Class B bondholders will receive $100 for sure, regardless of the quality of the securitized loan portfolio. Thus, no information communication by the bank is necessary. The question is: How much will North American have to pay the credit enhancer, assuming that the credit enhancement is competitively priced?

Step 4 To answer this question, suppose we label the two loans in the portfolio as 1 and 2. Then, there are four possible "states": (i) both loans 1 and 2 pay off $100 each (the probability of this is $0.9 \times 0.9 = 0.81$), (ii) loan 1 pays off $100 and loan 2 pays off $50 (the probability of this is $0.9 \times 0.1 = 0.09$), (iii) loan 1 pays off $50 and loan 2 pays off $100 (the probability of this is $0.1 \times 0.9 = 0.09$), and (iv) both loans pay off $50 each (the probability of this is $0.1 \times 0.1 = 0.01$). Now, in state (i), the credit enhancer has no liability since the total portfolio cash flow of $200 is enough to fully satisfy the claims of both classes of bondholders. In states (ii) and (iii), the total portfolio cash flow is $150, so that only $50 is available to pay off Class B bondholders after Class A bondholders are satisfied. In each of these states, the credit enhancer's liability is $50, and the total probability of these two states [that is, the probability that *either* state (ii) or state (iii) will occur] is $0.09 + 0.09 = 0.18$. In state (iv), the total portfolio cash flow is $100, so that the credit enhancer's liability is 100. Hence, the expected value of the credit enhancer's liability is $0.18 \times \$50 + 0.01 \times \$100 = \$10$. In a competitive market (with a zero discount rate and risk neutrality), this is what North American will have to pay for the credit enhancement. Thus, North American's net payoff will be

$100	+	$100	−	$10	=	$190.
market value of		market value of		credit		
Class A bonds		bonds with credit		enhancement		
		enhancement		fee		

The loan portfolio has been made "perfectly" liquid!

(3) Diversification of Funding Sources: Securitization provides originators with a way of raising financing from sources other than their traditional sources. For example, for a bank the traditional sources of funding are deposits, federal funds, subordinated debt, preferred stock and equity, all of which are essentially claims on the bank's entire asset portfolio. By segregating some assets into a pool to be securitized, the bank is able to diversify its funding sources beyond its traditional sources. This is important for any bank CFO.

(4) Enables Focus on Origination, Servicing, and Monitoring: You will recall from our discussions in Chapter 3 that the key economic functions of financial intermediaries are related to the resolution of informational problems. One advantage of securitization is that it enables the bank to focus on the origination, servicing, and monitoring of loans, three activities that the bank can generally perform more efficiently than others because of specialization. When a bank or S&L originates a loan, it provides a valuable screening service—the bank's willingness to make the loan tells other interested but less-informed parties something about the borrower that they did not know before. This can reduce the borrower's cost of credit from other (nonbank) sources. Loan servicing is a transactional service that the bank may be able to provide at lower cost because of its specialization in handling numerous other similar transactions. And of course, monitoring of borrowers is one of the fundamental intermediation services that banks provide. By securitizing its loans, a bank can focus on these three activities without actually funding the loans. Absent regulatory help, it is not clear that banks have any special advantage in funding anyway. More importantly, recent empirical research has shown that it is efficient for banks to specialize (not necessarily completely) either in nontraditional activities (like mortgage bankers who specialize in origination) or in traditional banking activities (which involve funding). That is, there are potential *diseconomies of scope* between traditional banking activities (like originating, monitoring, and funding) and nontraditional banking activities (like loan selling and buying).[23]

(5) Facilitates the Avoidance of "Adverse Selection" Costs: Typically a bank has a portfolio of assets with differing degrees of information sensitivity. Even if the market prices the portfolio correctly, it is quite possible that some of the assets are overvalued by the market and some are undervalued. Now consider a bank faced with the prospect of funding some new loans about which there is very little informational asymmetry, that is, the bank's assessment of the value of this new loan portfolio is roughly equivalent to that of the market. But suppose that the bank on the whole is undervalued. If this bank goes out and raises capital (that consists, in part, of uninsured deposits and equity) to fund these new loans in the conventional (nonsecuritization) manner, it will have to pay an "adverse selection cost" in the sense that the bank's cost of funding will be higher than it would be if investors had the same information about all of the bank's assets as the bank's managers themselves. This happens because investors who provide funding for these new loans are purchasing claims against the bank's *entire* asset portfolio.[24] That is, these factors may account for the bank's disadvantage in funding loans.

Instead of going the conventional route, suppose the bank decides to securitize these loans. Now investors who provide the necessary capital are purchasing claims only against the new assets. Thus, if there is little informational asymmetry about these assets, there will be virtually no "adverse selection" cost.[25] An example is

23. See Mester (1992).
24. See James (1988) for a similar point.
25. Ignore signaling complications created by the possibility that the market may revalue the bank's existing assets when the bank decides to securitize its new loans. This would occur because the market recognized that part of the bank's incentive to securitize the new assets comes from its own private knowledge about its existing assets.

provided by Gelco Corporation, a truck-leasing company with a BB credit rating from Standard & Poor's. Its commercial paper, backed by high-quality leases, was rated A-1. The firm saved about 80 basis points in borrowing costs by securitizing its lease receivables.[26] This illustrates how a firm can take advantage of its marginal cost of funds on all its assets.[27]

A key element of the benefit of avoiding adverse selection costs is that securitization through SPCs often achieves "bankruptcy remoteness" of the securitized assets from the borrowing firm. That is, the claims of lenders who provide financing to the SPC *cannot* subsequently be diluted by the claims of Debtor-in-Possession lenders to the sponsoring firm should the sponsor file for bankruptcy. However, this bankruptcy remoteness protection for the SPC investors is not perfect and can be undermined in legal proceedings. We would then expect that the greater the legal risk of bankruptcy remoteness being undermined by the courts, the higher will be the interest rates demanded by the SPC investors.[28]

(6) Avoidance of Intermediation Taxes: Because securitization permits a depository institution to raise funds directly from investors rather than from depositors, it helps the institution to avoid "intermediation taxes" (or regulatory taxes) like reserve and capital requirements and deposit insurance premiums. This benefit is obvious in the case of a pass-through because the relevant assets are removed from the institution's balance sheet, thereby eliminating the need to hold capital against those assets. Moreover, since the proceeds from the sale of the pass-throughs are not deposits, no reserve requirements or deposit-insurance premiums are involved.

There is a more subtle interaction between regulatory taxes and the institution's choice of which assets to securitize. Many regulatory taxes, like reserve requirements, have traditionally been *flat* taxes in that they do not depend on the riskiness of the asset involved, That is, there is "pooling" of these taxes across the spectrum of asset risks, and the actual taxes correspond to some sort of average.[29] Thus, the low-risk assets (which should have lower-than-average taxes) "subsidize" the high-risk assets (which should have higher-than-average taxes). A bank can lessen the impact of these taxes by securitizing its low-risk assets and leaving only the high-risk assets on its books. This way it would have higher-than-average risk assets on its books, but it would only pay taxes that correspond to average risk.

The Demand Side of Securitization: Investor's Perspective Having discussed why financial institutions and other firms might wish to securitize assets, we now turn to why investors might wish to hold these assets. One way for investors to invest in a bank's assets (loans, for example) is to purchase bank equity. Relative to that alternative, purchasing securitized claims offers a number of possible advantages, two of which are discussed below.

26. See Shapiro (1985).

27. The intuition here is similar to that for another fairly well-established practice known as "project financing," which involves a corporation establishing a legally distinct subsidiary to finance a new project. See the discussion in Chapter 7. Shah and Thakor (1987) provide a rationale for the high leverage ratios in project financing.

28. This is what Ayotte and Gaon (2006) find in their empirical tests.

29. This was suggested by Pavel (1986).

(1) Reduction in Market Incompleteness: These are two ways in which securitization helps to reduce market incompleteness. First, it improves the *quality* of assets that investors can hold and thereby increases the quality spectrum of available assets. Second, it provides a greater *variety* of cash-flow streams that investors can hold. Consider the quality issue first. The claims that an institution offers investors via securitization are often of higher credit quality than the institution itself. This is due to two reasons. First, as we have discussed earlier, there are incentives for the institution to securitize both lower risk and less information-sensitive assets. Second, credit enhancement improves the quality of the asset pool being securitized. Thus, it is not surprising that most of the asset-backed market is triple-A or double-A rated. Securitization provides investors with access to higher credit quality claims than available otherwise.

Now consider the cash-flow variety issue. Because securitization combines pooling, cash-flow partitioning, and credit enhancement, it does *not* produce merely a linear combination of existing payoff vectors (recall the discussion of market completeness in Chapter 1). Rather, it produces claims that were previously unavailable to investors through linear combinations of existing claims. Moreover, even in the case of claims that could have been "home-manufactured" by investors who were willing to combine available securities, securitization provides a less-expensive alternative in terms of transactions costs. For example, asset-backed securities have limited prepayment risk, so the effective maturity of the security is relatively insensitive to market yields. This means that for a given decline in yield, these "positive" convexity and limited prepayment features may not be available in the same configurations to investors in nonsecuritized alternatives. Thus, securitization helps to reduce financial market incompleteness.

(2) Liquidity: Because of the size of the asset-backed securities market and the active trading involved, investors are assured that they are buying a liquid claim. Securitization may be viewed as an alternative technology (to traditional funding) for producing liquidity.

What?: Securitization With Recourse Versus Deposits and Risk Sharing

The concept of recourse is key to understanding what securitization does for both the originator and the investor. When securitization is without recourse, the investor has a claim only against the pool of assets that have been securitized. He has no claim to any other assets of the originator. On the other hand, securitization with recourse closely resembles traditional balance sheet lending. The purchaser of a security with recourse has the option of trading the claim for a general bank claim like that of an uninsured depositor should the purchased asset default. If the bank fails, the investor has the option of keeping the securitized asset.[30]

30. This line of reasoning is based on Benveniste and Berger (1987). The model that appears a little later is also based on the analysis in Benveniste and Berger (1987).

In this subsection, we will discuss securitization with recourse.[31] It turns out that the securitization benefits are similar to those created by multiclass securities, in which sequential claims are issued against the same collateral pool. Basically, better risk sharing is achieved, since the most risk-averse investors can be sold the most senior claims. Under current law, banks are not allowed to issue multiclass or senior deposit claims against their balance sheet assets. That is, deposit claims cannot be prioritized. Hence, from the bank's perspective, a valuable opportunity to cater to "preference clienteles" among its potential depositor base is lost. However, securitization with recourse gives the bank an opportunity to profit from selectively catering to depositors with different degrees of risk aversion. Uninsured depositors can switch to a contract that gives them a *senior* claim on a part of the bank's asset portfolio (that is, the securitized asset). A loan-backed security with a bank guarantee attached (securitization with recourse) is like a large CD with the addition of a senior keeping the securitized security rather than waiting in line with other depositors to obtain his share of the bank's *other* assets. In a capital market in which people have different degrees of risk aversion, we would expect the more risk-averse investors to buy these securitized claims, and their less risk-averse cohorts to bear more risk. Thus, securitization achieves better risk sharing than the standard deposit contract. This intuition is formalized in the model developed in the box below.

The model: Consider a bank that needs to raise \$d to finance its first-period investment, which returns \$A(θ) at the end of the period, where θ is a possible future state of nature. Suppose that the bank must raise \$$d_i$ in FDIC-insured deposits, \$$d_u$ in uninsured deposits, and \$$d_e$ in equity where $d_i + d_u + d_e = d$. Insured deposits carry an interest rate of r_f (these deposits are riskless), which is the riskless rate. We assume that the FDIC and the bank's shareholders are risk neutral. There is a single, representative uninsured depositor who is risk averse and has a utility function, U(w), over wealth, which is increasing and concave, that is, $U'' > 0$, $U'' < 0$ (recall the discussion of risk-aversion in Chapter 1). The alternative to uninsured deposits is to invest in the risk-free asset, which carries an interest rate of r_f.

Let \$B be the repayment promised to the uninsured depositor and D(θ) as the state-contingent payment actually received by them at the end of the period. To induce them to invest in uninsured deposits, it must be true that

$$\sum_{\theta=1}^{N} p(\theta)U[D(\theta)] \geq U\{[1 + r_f]d_u\} \qquad [9.1]$$

where the possible future states are $\theta = 1, \ldots, N$, and p(θ) is the probability of state θ. The inequality in (9.1) says that the expected utility of the uninsured depositor from investing in these deposits can be no less than his utility from receiving a sure payoff by investing in the risk-free asset.

31. An example of this is the ABB discussed previously.

Now $[1 + r_f]d_i$ is the amount promised by the bank to the insured depositors, and $D(\theta)$ is the amount for uninsured depositors. If $A(\theta) \geq [1 + r_f] + B$, the bank is solvent and the insured depositors receive $[1 + r_f]d_i$ from the bank. If $A(\theta) < [1 + r_f] + B$, the bank fails. The insured depositors still receive $[1 + r_f]d_i$, but only a portion of it comes from the bank. The FDIC covers the rest. This is a situation in which the FDIC takes over the bank, pays off $[1 + r_f]d_i$ to the insured depositors, and then shares the remaining assets of the bank *proportionately* with the uninsured depositors.[1] The proportions are determined by the relative contributions of insured and uninsured deposits to the total *deposit base*. That is, whenever there is insolvency (that is, $A(\theta) < [1 + r_f]d_i + B$), the amount collected by the uninsured depositors is

$$D(\theta) = \left[\frac{B}{B + [1 + r_f]d_i}\right]A(\theta) \qquad [9.2]$$

and the amount collected by the FDIC is

$$F(\theta) = \left[\frac{[1 + r_f]d_i}{B + [1 + r_f]d_i}\right]A(\theta). \qquad [9.3]$$

We can now write down each party's payoff at the end of the first period. First, the insured depositors receive $[1 + r_f]d_i$ regardless of θ. Second, the uninsured depositors receive

$$D(\theta) = \begin{cases} B & \text{if } A(\theta) \geq B + [I + r_f]d_i \\ [B + [1 + r_f]d_i]A(\theta) & \text{otherwise.} \end{cases} \qquad [9.4]$$

The bank's shareholders receive (at the end of the period)

$$S(\theta) = \begin{cases} A(\theta) - B - [1 + r_f]d_i & \text{if } A(\theta) > B + [1 + r_f]d_i \\ 0 & \text{otherwise.} \end{cases} \qquad [9.5]$$

The FDIC receives (at the end of the period)

$$F(\theta) = \begin{cases} 0 & \text{if } A(\theta) \geq B + [1 + r_f]d_i \\ -[1 + r_f]d_i + \left[\dfrac{[1 + r_f]d_i}{B + [1 + r_f]d_i}\right]A(\theta) & \text{otherwise.} \end{cases} \qquad [9.6]$$

Note that the lower term in (9.6) applies when $A(\theta) < B + [1 + r_f]d_i$, so that

$$-[1 + r_f]d_i + \left[\frac{[1 + r_f]d_i}{B + [1 + r_f]d_i}\right]A(\theta) < 0.$$

Thus, the *end-of-period* cash flow to the FDIC is always zero or less.

At the start of the period, the bank's shareholders pay a deposit insurance premium to the FDIC. We assume that this premium is risk insensitive and fairly

(*Continued*)

priced. Let π denote this premium. To write down this premium, let us rank-order the states θ in increasing order of $A(\theta)$, so that $A(1) < A(2) < \ldots < A(N)$. Let $\theta = m$ be the state such that $A(\theta) \geq B + [1 + r_f]d_i$ for all $\theta > m$ and $A(\theta) < B + [1 + r_f]d_i$ for all $\theta \leq m$. Then, the expected value of (9.6) is

$$\pi = \frac{1}{[1+r_f]} \sum_{\theta=1}^{m} p(\theta)\left\{[1+r_f]d_i - \left[\frac{[1+r_f]d_i}{B+[1+r_f]d_i}\right]A(\theta)\right\}. \qquad [9.7]$$

That is, the deposit insurance premium is equal to the discounted present value of the FDIC's liability.

The question is: What is the NPV of the bank's shareholders' investment? This is seen to be

$$\text{NPV} = \frac{1}{[1+r_f]} E[S(\theta)] - \pi - d_e \qquad [9.8]$$

where E(•) means "expected value" and $S(\theta)$ is given by (9.5). Now note how B and d_u are linked. Using (9.1) and (9.4) we see that they are linked as follows.

$$\sum_{\theta=1}^{m} p(\theta)U\left\{\left[\frac{B}{B+[1+r_f]d_i}\right]A(\theta)\right\} + \sum_{\theta=m+1}^{N} p(\theta)U(B) = U\{[1+r_f]d_u\}. \quad [9.9]$$

The left-hand side (LHS) of (9.9) is the uninsured depositor's expected utility from investing in that deposit; this is obtained directly from (9.4). We are using (9.1) as an equality here because the bank, whose objective is to maximize the wealth of its shareholders, will pay the *minimum* amount required to attract funds from the uninsured depositor.

It is clear now that the higher d_u is, the higher B will have to be to satisfy (9.9). Raising B has two effects. One is that it increases the amount collected by the uninsured depositor when the bank is solvent. The other is that it increases the uninsured depositor's proportional share of the bank's assets in insolvency.

For any fixed d_u, B must also increase as the uninsured depositor becomes more risk averse. This is because he demands a higher risk premium, or a higher expected value for his risky payoff. Of course, as B increases, $S(\theta)$ decreases, so that the bank's shareholders become worse off as the uninsured depositor becomes more risk averse [see (9.8)]. This provides a strong motivation for securitization with recourse because of the possibility of reducing the risk borne by the uninsured depositor. We will now see how this is achieved.

To incorporate securitization into this model, let us partition the bank's assets into two portfolios—a balance sheet portfolio, A_b, and an off-balance sheet portfolio, A_o. The loans in A_b are funded with insured deposits and equity, which add up to $d_i + d_e$ dollars. The loans in A_o are funded with securitized bonds that fetch the bank \$$d_u$. These two portfolios combined give the same payoff as before, that is,

$$A(\theta) = A_b(\theta) + A_0(\theta). \quad [9.10]$$

We assume that the bank continues to service the loans' securities. In doing so, the bank directs receipts of loan payments to the depositors. The key feature of this new arrangement, however, is that the receipts from $A_0(\theta)$ are committed to repay the securitized bonds first. Only after these bonds are fully paid off can the revenues be directed elsewhere. That is, the securitized bondholder has a senior claim on the payment stream $A_0(\theta)$. Adding the option of securitized bond gives the holder additional protection by allowing him to change his claim to that of a balance sheet liability if the revenues from $A_0(\theta)$ are inadequate. In this case, he exercises his claim just like an uninsured depositor, as $A_0(\theta)$ and $A_b(\theta)$ are pooled together. Thus, when the securitized bondholder exercises his recourse option, he is limited to a prorata claim on the total portfolio $A(\theta)$. Let B^* be the amount promised to the securitized bond.

The securitized bondholder thus receives

$$\begin{aligned} &B^* \quad \text{if } A(\theta) \geq B + [1 + r_f]d_i \\ &B^* \quad \text{if } A(\theta) \geq B + [1 + r_f]d_i \text{ and } A_0(\theta) \geq B \\ &D^*(\theta) = \left[\frac{B^*}{B^* + [1 + r_f]d_i}\right] A(\theta) \quad \text{otherwise.} \end{aligned}$$

This payoff structure is illustrated in the figure below. We have defined $\theta = n$ as the state in which $B^* = A_0(\theta)$.

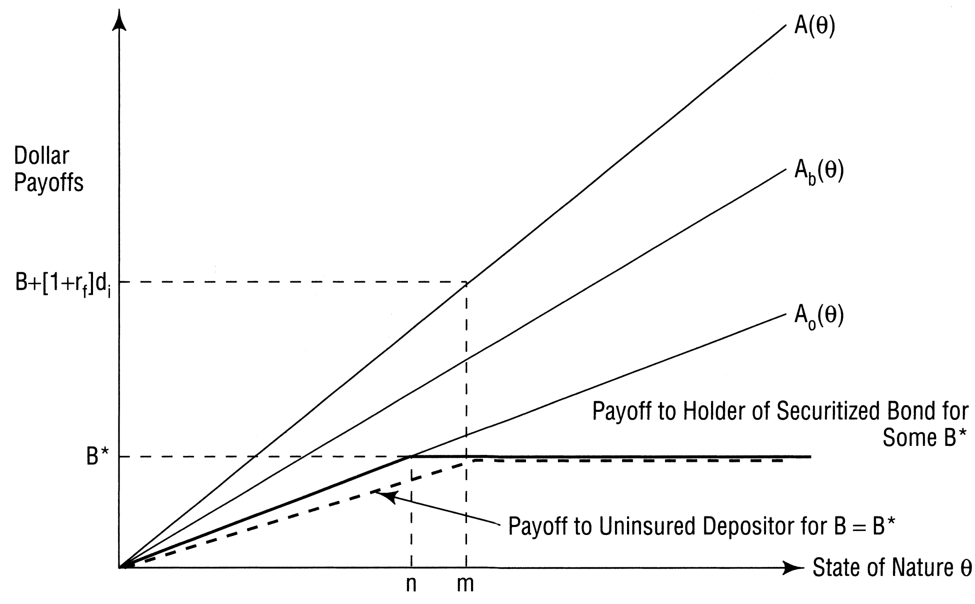


FIGURE 9.9 State-Contingent Payoff to Securitized Bondholder

(*Continued*)

In this figure, we have assumed, for comparability, that $B^* = B$, that is, the promised repayment amount is the same for uninsured deposits and the securitized bond. In that case it is clear that securitization gives the uninsured depositor a higher payoff for the same repayment amount. This higher amount comes from the additional-option feature embedded in the securitization contract. In the context of our model, what is held fixed across the securitization and no-securitization alternatives is not the promised repayment, but the initial amount raised from uninsured creditors, which is d_u. When d_u is held fixed, we have $B^* < B$.

Rather than formally prove that securitization with recourse improves the wealth of the bank's shareholders, we will provide a numerical example to make the point. The basic intuition is simply that both the FDIC and the bank's shareholder are risk neutral and hence "better able" to absorb risk than the risk-averse uninsured depositor. Securitization with recourse transfers some risk away from the uninsured depositor to these parties, thereby improving risk sharing. The effect is to reduce the amount that the bank's shareholders must promise to repay the uninsured creditor, which increases the expected value of the shareholders' claim. It also increases the riskiness of their claim, but they do not care about that because they are risk neutral.

1. Under current law, senior claims are not permitted on the balance sheets of banks. Thus, when a bank fails, the FDIC pays the insured depositors in full and shares the assets of the bank proportionately with the uninsured claimants in a manner similar to that indicated in the development below.

Example 9.3 The North American Bank needs to raise \$50 in financing at the beginning of the period to finance an investment that will yield a random payoff at the end of the period. This payoff has the following probability distribution.

State θ	**1**	**2**	**3**	**4**
Payoff $A(\theta)$	\$0	\$50	\$100	\$150
Probability $p(\theta)$	0.25	0.25	0.25	0.25

The single-period riskless interest rate is 0.10. The bank must raise \$30 in FDIC-insured deposits, \$15 in uninsured deposits (with or without securitization) and \$5 in equity. North American's shareholders and the FDIC are risk neutral. The uninsured depositor is risk averse with utility $U(w) = \sqrt{w}$ for wealth $w \geq 0$. Compute the NPV to North American's shareholders when: (i) they finance conventionally using deposits and (ii) when they issue a securitized bond that has a preferred claim on a specially designated asset portfolio that represents 60 percent of the payoff on $A(\theta)$ in any state.

Solution We solve this problem in six steps. First, we consider conventional deposit financing and solve for B, the repayment that must be promised to the uninsured depositors in order to raise \$15 from them. We do this by using (9.9). Second, we solve for the insurance premium, π, using (9.7). Third, we compute the NPV to North American's shareholders. Fourth, we consider the alternative of issuing a securitized bond and solve for B^*, the amount promised to the securitized bond. Fifth, we solve for the insurance premium in this case. Finally, in step 6 we solve for the NPV accruing to North American's shareholders with securitization, which is higher than the NPV accruing to them with conventional deposit financing.

Step 1 Let us consider (i) first. Note that $d_i = 30$, $d_u = 15$, $d_e = 5$, and $r_f = 0.10$. Since the claim of insured depositors is riskless, the interest rate on these deposits should be the riskless rate, 10 percent. Thus, North American's repayment obligation to these depositors is $d_i[1 + r_f] = 30[1.1] = \$33$. The bank's total repayment obligation to its creditors is $33 + B$. To solve for B, we need to *conjecture* about its value first, so that we can identify the states in which North American is solvent and those in which it is not. Suppose we conjecture that $50 < 33 + B < 100$. Then North American will be insolvent in states 1 and 2, and solvent in states 3 and 4. We now solve for B using (9.9):

$$\underbrace{0.25\sqrt{\left[\frac{B}{B+33}\right]\times 50}}_{\substack{\text{expected utility} \\ \text{payoffs in states 1 and 2}}} + \underbrace{0.5\sqrt{B}}_{\substack{\text{expected utility} \\ \text{payoffs in states 3 and 4}}} = \sqrt{1.1 \times 15}.$$

Solving this equation gives $B = \$32$, approximately. This means our conjecture is valid and that $B + 33 = \$65$.

Step 2 Next we solve for the insurance premium using (9.7).

$$\pi = \frac{1}{1.1}\left\{0.25 \times 33 + 0.25\left[33 - \left(\frac{33}{65} \times 50\right)\right]\right\} = \$9.23.$$

Step 3 Further, since North American's shareholders receive something only in the solvency states (3 and 4), the expected value of the payoff for shareholders is

$$E[S(\theta)] = 0.25[100 - 65] + 0.25[150 - 65] = \$30.$$

Thus, the NPV of North American's shareholders is given by (9.8) as

$$NPV = \frac{30}{1.1} - 9.23 - 5 = \$13.04.$$

(*Continued*)

Step 4 Now consider (ii). With securitization, the payoff stream of the total asset portfolio is split up as follows.

TABLE 9.2 Payoff Distributions for Portfolios

State	1	2	3	4
Payoff $A(\theta)$	0	50	100	150
Payoff to securitized bond$_f A_0(\theta)$	0	30	60	90
Payoff from on-balance sheet asset$_f A_b(\theta)$	0	20	40	60

Now, suppose we conjecture that $50 < B^* + 33 < 100$ and that $B^* \leq \$30$. Then, state 1 is the only state in which the holder of the securitized bond gets less than the amount promised. Thus, B^* is obtained by solving the following equation.

$$\underbrace{0.75}_{\substack{\uparrow \\ \text{cumulative} \\ \text{probability} \\ \text{of states} \\ \text{2, 3, and 4.}}} \times \sqrt{B^*} = \sqrt{1.1 \times 15} = 4.062,$$

which yields $B^* = \$29.33$.

Step 5 We can now solve for the deposit insurance premium, which is

$$\pi = \frac{1}{[1 + r_f]}\left\{ p(1)[1 + r_f]d_i + p(2)\left[[1 + r_f]d_i - \left\{A_b(2) + A_o(2) - B^*\right\}\right]\right\}.$$

Note that the FDIC is liable for payments only in states 1 and 2. In state 1, it is liable for the entire repayment promised to the insured depositors, whereas in state 2, it is liable for that amount minus what it collects on the balance sheet asset, $A_b(2)$, and whatever is left over on the off-balance sheet asset after the securitized bondholder is paid, $A_o(2) - B^*$. Thus,

$$\begin{aligned}\pi &= \frac{1}{1.1}\left\{0.25 \times 33 + 0.25\left[33 - \left\{20 + (30 - 29.33)\right\}\right]\right\} \\ &= \$10.30.\end{aligned}$$

Step 6 Now, $B^* + [1 + r_f]d_i = 29.33 + 33 = \$62.33$. Thus, $E[S(\theta)] = 0.25[100 - 62.33] + 0.25[150 - 62.33] = \$31.34$.

This means

$$\text{NPV} = \frac{31.34}{1.1} - 10.30 - 5 = \$13.19.$$

Thus, in this example securitization results in a 1.16 percent increase in the NPV accruing to the North American Bank's shareholders, and this gain is due to the differing risk preferences of uninsured depositors and the banks and their insurer.

How Much Securitization?

Given the proliferation in asset-backed securities and the spread of securitization to even intangible assets, an obvious question might be: Can and should everything be securitized? The answer is no, and there are three main factors that explain why there are still assets that have not been securitized.[32] These are discussed below.

(1) Ease of Standardization: First, for an asset to be profitably securitized, it should be relatively easy to "standardize." That is, its contract features should make it a component of a relatively homogeneous portfolio of other similar assets. Mortgage loans are an excellent example. Mortgage contracts are standardized and their cash-flow patterns are, on average, quite predictable. This was not always the case. Indeed, the possibility of securitization prompted the standardization. Loans with special contract features (such as, HLT loans and some types of LBO loans secured partly by personal collateral of managers) are difficult to standardize and hence would be difficult to securitize, although we have already witnessed the securitization of LBO loans. The reason for the desirability of contractual homogeneity is related to transactions and information-processing costs. When a large number of individual loans are pooled, the investor only has to evaluate the portfolio return. This creates both "informational diversification" and "statistical risk diversification," so that the portfolio cash flows are more predictable and less sensitive to the peculiarities of the individual assets. This eases the originator's/issuer's problem of designing specific securities based on the total portfolio cash flow. Moreover, it makes it easier for investors to evaluate the values of the securitized assets. This attracts a larger number of investors and improves the liquidity of the market for the securitized assets.

This is not to say that more heterogeneous and information-sensitive assets, which are more difficult to standardize, cannot be securitized. Rather, such assets would require greater credit enhancements to be securitized; the limiting case is that of "full-recourse" securitization. When the assets that comprise the portfolio are very dissimilar and difficult to standardize, the portfolio cash flows become quite sensitive to the actual choice of securities that make up the portfolio. More credit enhancement would be needed for such a portfolio, and beyond some point, it may not be worthwhile for the originator to purchase the required credit enhancement. Thus, standardization can be viewed as conserving the originator's capital in that a high credit rating can be obtained for the portfolio with *less* credit enhancement. Without the appropriate credit enhancement, the rating agencies and investors have to study individual securities in the portfolio more. To see this more clearly, consider the following example.

32. See Caouette (1990).

Example 9.4 Suppose there are three possible assets from which the North American Bank can choose two to securitize. Call these assets a, b, and c. The assets are quite similar and their cash-flow distributions are as described below. The probability of "success" for asset i is p_i. Compute the probability distributions of the various portfolio combinations. How important would it be for an investor to know precisely which two assets were in the securitized portfolio? Would your answer change if asset c were replaced by asset d, which has a cash flow that is uniformly distributed over [0, 1000]?

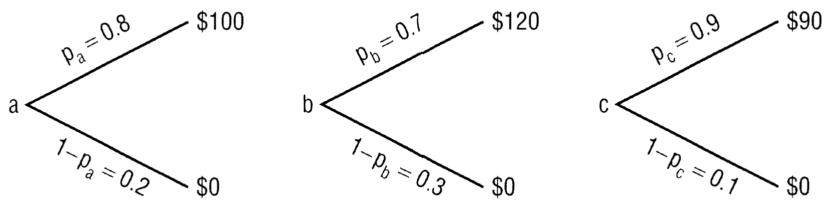


FIGURE 9.10 Probability Distributions of Assets a, b, and c

Solution We will solve this problem in two steps. First, we will compute probability distributions for the three possible portfolio combinations—ab, ac, and bc—and conclude that they are pretty similar. Second, we will show that replacing c by d would make a significant difference in that the investor will now need to know the portfolio composition.

Step 1 Now, if North American were to choose assets a and b, the total portfolio cash-flow distribution would be: 0 with probability 0.06 (which is $[1 - p_a] \times [1 - p_b]$), 220 with probability 0.56 (which is $p_a \times p_b$), 100 with probability 0.24 (which is $p_a \times [1 - p_b]$), and 120 with probability 0.14 (which is $[1 - p_a] \times p_b$). The expected value would be 164. If the issuer were to assemble the portfolio with assets a and c, the probability distribution would be: 0 with probability 0.03, 210 with probability 0.27. The expected value would be 161. Similarly, if the portfolio consisted of b and c, the probability distribution would be: 0 with probability 0.03, 210 with probability 0.63, 120 with probability 0.07, and 90 with probability 0.27. The expected value would be 165. The portfolio payoffs are summarized below.

TABLE 9.3 Probability Distributions of Various Portfolios

State/Portfolio	Low Cash Flow and Probability	Medium Cash Flow and Probability	Above-Medium Cash Flow and Probability	High Cash Flow and Probability	Expected Value
a and b	0 w.p. 0.06	100 w.p. 0.24	120 w.p.0.14	220 w.p.0.56	164
a and c	0 w.p. 0.02	90 w.p. 0.18	100 w.p.0.08	190 w.p.0.72	161
b and c	0 w.p. 0.03	90 w.p. 0.27	120 w.p.0.07	210 w.p.0.63	165

Clearly, if you were an investor deciding whether to buy a piece of the portfolio, it would not be terribly important to know the precise composition of that portfolio. All the portfolios have similar expected values and each portfolio has a low probability of a low cash-flow realization, a high probability of a high cash-flow realization and intermediate probabilities for the medium and above-medium cash flows.

Step 2 Now imagine that c is replaced by asset d, which has a cash flow that is uniformly distributed over [0,1000]. It is now easy to see that it will be quite important for the investor to know which assets are in the portfolio. For example, a combination of a and d has an expected value of $580 (the expected value of a is $80 and the expected value of d is $500), whereas a combination of a and b has an expected value of $164. And if the portfolio cash flows are partitioned, it will be even more important for the junior claimants to know the portfolio composition. This is why homogeneous pools of assets are easier to securitize.

(2) Extent of Private Information: Another important consideration is the extent of private information about the asset. If the loan originator has substantial information about the loan that others do not have, then information communication costs might deter securitization. This point is related to homogeneity in that information-sensitive assets are more difficult to standardize. But, even if the loan portfolio is homogeneous, it may be costly to securitize if each loan in the portfolio is steeped in private information possessed by the originator. Truly "opaque" assets are usually difficult to sell at anywhere close to their true value. Consequently, steep discounts may be needed to entice investors to buy assets they do not fully understand. It may not benefit the originator to securitize such assets. For example, it would be difficult to securitize and sell in the United States a portfolio of consumer loans made by a local bank in Nigeria, particularly if reliable statistics on historical repayment patterns were unavailable. By contrast, U.S. credit-card receivables are relatively easy to securitize in the United States. There is only cursory initial screening of credit-card applicants, so there is not much that the lender knows that others do not know. Moreover, the contract itself is fairly standardized, and repayment patterns of *portfolios* of credit cards are quite predictable. A similar argument holds for consumer mortgages. Screening procedures for determining who gets a mortgage loan are standardized, so that once a person is given a mortgage contract, she falls in a pool about which the original lender knows little more than the rating agencies and investors.

One important implication of how private information affects assets chosen for securitization is that the quality of the assets on the balance sheets of banks may deteriorate. The assets that are securitized are typically more liquid. Assets that stay on the books are likely to be less liquid and have other problems as well.[33] The bank may prefer to keep these lower-quality assets on its books because it may be able to sell them only at steep discounts relative to their "true" values, which may be privately known only to the bank. Since there is often an inverse relationship between liquidity and risk, the bank's portfolio risk *may* increase due to securitization. Mitigating this concern is the fact that banks can diversify by buying securitized claims against loans originated by others.

33. Greenbaum and Thakor (1987) show formally that, under certain conditions, banks will securitize higher-quality (lower default risk) assets and retain on their balance sheets assets of lower quality.

Of course, as information technology advances, the costs of processing and communicating financial information decline. This makes securitization less costly. The obvious implication is that securitization can be expected to grow in scope and volume.

(3) Moral Hazard: With traditional lending, the original lender combines origination, underwriting, funding, and servicing of the loan. The lender then has an incentive to monitor the loan. As we saw in Chapter 3, monitoring is an important activity of banks. With securitization, however, origination and funding are separated. This weakens the originator's incentive to monitor the loans in the securitized portfolio.[34] The reason is that monitoring is costly to the originator, and the benefits of the monitoring—an improvement in the cash flows from the securitized portfolio—accrue to the investors who have purchased the securities, not to the originator. In this context, the traditional bank can be seen as a solution to the moral hazard that can accompany loan decomposition. An obvious solution is to shift some of the credit risk back to the originator by employing securitization with partial recourse. This places the exposure with the party responsible for the monitoring, and hence reduces moral hazard.

However, recourse raises other accounting/regulatory problems. Recall that one strong incentive to securitize comes from capital and reserve requirements, deposit insurance premiums, and other costs, which add an estimated 125 basis points to the funding costs of deposit-takers. If loans are securitized with recourse, they are not usually removed from the books, and hence none of the regulatory costs are avoided. However, the originator can utilize alternatives to recourse as a way to deal with moral hazard without keeping the loans on its books. Senior/subordinated structures, overcollateralization, and third-party guarantees can all provide credit enhancement to attenuate moral hazard.

Credit enhancement deals with moral hazard in two ways. First, it directly improves the credit quality of the securitized asset, so that investors are less affected by the quality of underwriting and monitoring provided by the originator. Second, it creates an incentive for the credit enhancer to *monitor the originator* to ensure that the originator is underwriting and monitoring the loans in the securitized portfolio. We can expect the credit enhancer to be specialized in monitoring and thus monitor the originator more efficiently than individual investors can. Moreover, just as a bank saves on monitoring costs by centralizing the monitoring activity and thereby avoiding *duplication of efforts* (Chapter 3), a credit enhancer can save on monitoring costs that would otherwise be greater because of duplicated monitoring by individual investors. There are, however, natural limits to the gains from credit enhancement since the marginal effectiveness of the credit enhancer's monitoring will expectedly decline with more monitoring (diminishing returns to scale). This will be reflected in the fee charged to the originator by the credit enhancer. Depending on the nature of the asset (in particular, the sensitivity of its cash flow to monitoring by the originator) and the level of credit enhancement sought, a point may be reached beyond which further credit enhancement is not justified from the originator's standpoint. And there may well be assets that are better for the originator to fund with deposits than to securitize and credit enhance up to the optimal limit, because of moral hazard.

Some types of commercial and industrial (C&I) loans are unlikely candidates for securitization, although we anticipate this to be a shrinking list. Loans whose values are

34. This reasoning appears in Greenbaum (1987) and Gorton and Pennachi (1993). Mester (1992) provides empirical evidence that it is less costly for a bank to monitor a loan it has originated itself than to monitor a loan it has purchased.

highly dependent on lender monitoring are usually subject to a great deal of moral hazard, as we have seen in Chapters 5 and 6. Such loans are difficult to securitize because the required credit enhancement would be too costly. Put differently, unpredictability in the quality of monitoring provided by the originating lender may lead to a lot of unpredictability in the cash flows generated by such loans. We would expect, on moral hazard grounds, that such loans would be securitized infrequently. This is what we observe. However, despite this, we expect few, if any, C&I loans to not be securitized in the future.

Strategic Issues for a Financial Institution Involved in Securitization

Securitization is technology for liquefying claims and diversifying funding sources management by banks and other financial institutions. Its enormous growth is only one of the indicators of an ongoing revolution of ideas in the capital markets.[35]

Securitization as a Balance Sheet Management Tool

As already seen, securitization can be used by a financial institution to manage a variety of risks. These are:

- Interest rate risk,
- Prepayment risk,
- Credit risk,
- Liquidity risk.

Most of the loan sales and loan participations involve sales without explicit recourse to the originating lender, so that all the risks listed above are sold off together. In the ABS market, recourse is much more common, so that only some of the risks are disposed of by the seller. The reason for this difference is not hard to see. Loan sales and participations involve banks dealing with each other or with other financial institutions that are quite capable of assessing the risks involved and ensuring adequate monitoring.[36] Recourse, with its associated regulatory costs, is therefore avoided. But securitization involves a financial institution dealing with smaller investors who must be reassured through recourse that adequate monitoring will be provided.

The management implication is that the financial institution must carefully balance the gains from laying off risk through securitization (without recourse) against the cost of securitizing. On the one hand, deposit funding, which is an alternative to securitization, has an interest cost plus a regulatory cost. On the other hand, securitization involves a fixed cost (legal costs plus distribution costs) plus a credit-enhancement cost (which is a function of the heterogeneity and information sensitivity of the asset pool). These costs of alternative funding modes will determine the institution's choices about which assets to securitize and how much

35. See Lelogeais (1990) and Bevier (1990).

36. In the case of loan sales too, the seller sometimes sells only a part of the loan, retaining a portion on its books. This may be viewed as a substitute for credit enhancement.

recourse to offer the buyers. As we have seen earlier, assets that are very "opaque" may be hard to securitize. The institution may prefer to either involve other lenders in a participation arrangement or fund it entirely by itself. In other words, balance sheet funding will be chosen when the financial institution has a comparative advantage in "piercing the veil" of an "opaque" asset but finds it difficult to explain its value to investors.[37]

Securitization as a Pricing Tool

With easier entry into banking, over the past several years banks have had to come up with ways to cope with greater price competition. While banks must be competitive, they also must make sure that they are pricing at profitable levels. Some banks attempt to do this by setting a minimum interest rate at which they will originate loans. They call this the "minimum buy rate" and set it at their estimated cost of funding those loans, plus a spread that provides a minimum acceptable return.[38] Unfortunately, banks often cannot determine precisely the real funding cost of any particular asset. Many use rough approximations based on their *average* overall cost of funds.

Securitization can help the bank improve the accuracy of its pricing. For example, if a bank is continually putting originated assets into a securitized pool, it can base its pricing directly on the funding costs for that asset. This will ensure consistent market-driven pricing and guarantee that the bank is always originating assets at a profitable spread. Moreover, since securitization helps the originating bank to avoid the costs of reserve and capital requirements, it may provide a competitive advantage over nonsecuritizing bank competitors. This permits the bank to either lower its prices and increase market share, or earn higher profit margins at current prices.

Securitization as a Strategic Weapon for Market Penetration and Diversification

Market Penetration: After having conducted a thorough analysis of its comparative advantage in specific markets, suppose your bank decides that its optimal strategy is to be a price leader in a given market and to expand market share. However, you are worried that there is substantial risk in that rapidly changing market conditions could mean that net interest margin (the difference between the loan interest rate and the bank's cost of funding that loan) fluctuates randomly, sometimes falling below the minimum required return on allocated capital. For strategic reasons, continuous presence in the market is necessary for your bank. The randomness in your future interest margin means, however, that such strategically motivated continuity may prove quite costly. It could force your bank to suffer periodic losses in order to retain its foothold in the market. This is a form of underinvestment problem in that you may be forced to forgo investment because of short-run costs, even though continued presence may be best for you in the long run.

37. See Lelogeais (1990).
38. See Bevier (1990).

An intelligently designed securitization program can be an excellent strategic weapon in overcoming this impediment. Suppose your bank has targeted the market for high-quality, low-margin auto loans. Then you could establish participation in a commercial paper program into which these loans can be periodically placed. Pricing on the auto loans can be adjusted daily to be in tune with market conditions. You can determine the margin on any loan at the time of origination, and compute the profitability of funding the loan with deposits as well as funding it through your participation in a commercial paper program. If the loan is unprofitable with either the deposit funding or the securitization alternative, then you may want to let the lending opportunity pass. If it is most profitable to fund the loan with deposits, you could mark it to be held on your bank's balance sheet. But suppose that this loan is unprofitable to fund with deposits. Then, rather than pass up the lending opportunity, you could examine whether it is profitable to securitize the loan. There will be instances in which securitization (at the commercial paper cost) is profitable for you even when balance sheet lending is not. Thus, the bank can ensure greater continuity of presence in its strategically chosen market segment by resorting to securitization. In essence, this is an argument for maintaining numerous funding sources even when there is a fixed cost to doing so. Note that such a strategy could also be motivated for liquidity reasons.

Diversification: As we saw in Chapter 5, asset portfolio diversification can significantly reduce a bank's risk of ruin. This is hardly a new insight. Virtually all bankers are aware of it. Yet there seems to be a trend toward increased *concentration* in loan portfolios.[39] This is particularly true for commercial and industrial loan portfolios. The reason is that increasing competition has forced banks to carve out competitive niches in order to maintain profit margins. Banks have specialized their lending in markets where they have a comparative advantage. And this leads quite naturally to geographic or industry concentration, or both. A conflict arises, therefore, between *specialization* and *diversification*.

Bankers initially responded to this conflict by engaging in international lending and by expanding through nationwide networks of loan production offices that were intended to overcome the shackles of state banking laws. For many banks, however, the results were disappointing. Losses resulted from an inadequate knowledge of the new credit environment and from the increased marketing expenditures needed to develop new relationships. The reaction to these unsuccessful diversification attempts was a retreat to familiar regional/industry market segments.

Securitization can rescue loan portfolios from underdiversification by permitting banks to overcome geographic limitations. Banks can originate and service loans to borrowers about whom they know more than others. The rewards for their superior knowledge will be reflected in the higher spreads and profit margins. But there is no need to sacrifice diversification by also funding these loans. The originating bank can securitize some of these loans and thereby free up capital to pursue other lending opportunities and diversify. Moreover, the bank also can purchase loans securitized by banks that specialize in *other* industries and geographical areas. This too will facilitate diversification.

39. See Haidorfer (1990). Boyd and Smith (1993) provide a related perspective on securitization. They argue that a lender will be more efficient in monitoring borrowers in its own area than in monitoring borrowers in another area. Thus, a bank will often find it optimal to diversify by purchasing some assets originated and securitized by other banks rather than originating and monitoring all its assets.

From the standpoint of the bank's overall strategy, securitization opens doors that were previously closed. A bank can aggressively pursue new markets without having to plan for periodic losses in the future just to retain its foothold in those markets, as well as specialize in industries and regions without risking under diversification.

Strategic Decisions and Securitization Structures

As the regulatory and competitive environment for banks further evolves, securitization will have to continue to adapt as it has adapted until now. Corporate ingenuity has thus far come up with innovative securitization structures that permit the implementation of the tactical and strategic goals of banks, and has resulted in an increasing number of tangible and intangible assets being securitized. We expect this trend to continue.

Comparison of Loan Sales and Loan Securitization

A loan sale, which we considered in Chapter 7, is similar to asset securitization with three basic differences. First, a loan sale merely transfers a part of the ownership of the loan portfolio from the originator to others, whereas securitization alters patterns of cash flows and other asset properties. Thus, a loan sale is a brokerage function, whereas securitization involves qualitative asset transformation (recall Chapter 2). Second, with securitization, claims against the asset are sold as securities in the capital market so that the original loans are converted into market-traded securities, whereas with a loan sale the asset is merely transferred from one bank to the other without material qualitative transformation. Third, most loan sales are made without explicit recourse to the seller. Moreover, unlike securitization, there are usually no guarantees, insurance, or any other type of explicit credit enhancement, although the portion of the loan retained by the originator functions somewhat like credit enhancement in attenuating moral hazard. Thus, a loan sale usually removes the loan permanently from the seller's balance sheet.

One striking difference between loan sales and securitization is that with a loan sale to another bank, the asset stays within the banking industry, whereas with securitization, it is converted into a capital market investment. However, loan sales and securitization provide a bank with similar advantages in terms of strategic choices. The bank's ability to specialize in originating, servicing, and monitoring loans to borrowers in specific regions and from specific industries can be put to profitable use without compromising diversification objectives, when the bank designs and implements a comprehensive loan sales and asset-securitization program.

Conclusion

Securitization and the Evolution of Banking

Securitization, which involves the qualitative (asset) transformation of loans into traded securities, may well be the most important engine of reform in our financial system. The trend toward securitization is likely to continue in the absence of a substantial decrease in the relative cost of funding for banks and thrifts. Funding advantages of the past were based on deposit interest rate ceilings, discounting and advances, deposit insurance and the tax system, and regulated entry into banking. It is unlikely that these advantages will be revived. Thus, securitization is likely to grow, as long as workable solutions to the moral hazard problem can be developed. If these solutions involve some form of recourse on the part of the originator, then they will need to satisfy accountants, regulators, and the contracting parties.

The securitization of thrift assets is much more advanced than that of commercial bank assets. This is mainly because Fannie Mae, Freddie Mac, and Ginnie Mae resolved the moral hazard problem of separating underwriting from funding by simply accepting the credit risks themselves in exchange for a fee. Only recently have some thrifts come to realize that these agencies, so instrumental in liquefying thrift mortgage portfolios and in providing subsidized guarantor services, may also in the long run obviate the need for thrifts. This is at the heart of the proposal to privatize Fannie and Freddie. If the role of thrifts in funding is diminished, the originating/underwriting role becomes critical. If a private party shares part of the credit risk on assets originated by the thrift, then it becomes important to this party how the thrift evaluates and processes the credit risk at the time of origination. And, if the thrift specializes in these intermediation functions, it can be expected to earn some rents. But when governmental agencies provide the credit-risk guarantee on streamlined underwriting standards, the last of the financial intermediary's roles may be jeopardized. Thus, these agencies, which were created to *complement* and assist thrifts in their service of housing, become a Trojan horse!

For commercial banks, there is no quasi-governmental promoter of a secondary market in loans. Although moral hazard inhibits the emergence of securitization for some bank assets, we can expect banks to develop workable risk-sharing contracts to overcome the problem. Such an advance is likely to open the floodgates to the securitization of hitherto unsecuritized assets, which would merely be a continuation of a trend we have witnessed in the past decade.

With the growth of securitization, some banks may simply become specialists in evaluating credit risks and monitoring borrowers. This would create the impetus for a new banking system. A new payments system might emerge around the securities markets, mutual funds, and credit cards.

In the longer term, if the total separation of origination and funding occurs, banks' link with the payments system may be weakened. This may call into question the need for bank regulation itself; after all, a common rationale for regulation is that the safety and soundness of the banking system are desirable because banks affect the money supply and hence the transmission of monetary policy.

Case Study Lone Star Bank

Introduction

Lone Star Bank is a relatively small but rapidly growing regional bank based in Palo Alto, California. In recent years the bank has specialized in loans to small personal computer manufacturers. Historically, these customers have been mostly located in and around the Palo Alto and surrounding "Silicon Valley" area, but in the past few years, the bank has pursued similar business in other high-tech growth communities.

Although the bank is not as large as some of its competitors, it has generated a strong reputation as a bank that understands the computer industry.

John Langston, Chief Executive Officer (CEO) of Lone Star, believes that growth in the personal computer industry will continue to be substantial. The bank would like to parlay its current reputation and expertise into a much larger presence, but faces two major constraints. First, keeping up with the fast growth of the industry will require the bank to deal with difficult funding issues, since the base of deposits available to the bank is limited by a variety of regulations and competitive considerations. In particular, leverage-ratio constraints would require the bank to raise additional capital to support a larger deposit base, and this is considered a costly alternative. Second, the bank must maintain a fairly continuous presence in the market for strategic reasons. A loss in market share could allow competitors to develop equally strong reputations for understanding the business.

With these considerations in mind, Langston calls Lana Tanner and Hugh Akston, executive vice presidents, into his office to discuss his concerns about the bank's loan pricing policies.

The Initial Meeting

Langston: You both know our situation. Since we've been doing these computer company loans, we've had great success. We saw an opportunity that other banks didn't understand. But I'm concerned about a couple of things. First, our business loan portfolio is getting to be really heavily concentrated in the computer industry. (*See Exhibit A*.) Second, I'm worried that our success in lending to this segment might mean that we aren't pricing these loans correctly. Our return on assets has been slipping a bit lately (*see Exhibit B*), and I think we may be pricing our loans too cheaply.

Tanner: Well, the obvious thing to do would be to try to diversify by going into some other markets. But we can't forget the fiasco we had a couple of years back, when we thought we would expand heavily into real estate development loans. (*See Exhibit C*.)

Langston: (*He groans*.) Ugh, that's for sure. We took a hit on that one. I think we've learned a lesson there. We should stick to what we are good at.

Akston: Well, the next thing to look at would be the sale of some of our assets. By transferring some of the funding of loans elsewhere, we would concentrate on what we do best: identifying and monitoring successful firms. That way, we could stay in this industry segment while simultaneously diversifying our portfolio. I've talked to some other banks about doing some loan sales.

Langston: Great! Does that look promising?

Akston: Unfortunately, it really doesn't. Most of the companies we deal with are fairly small, as you know. The fact that all of these companies are producing a product that is something like a commodity to the end-user means that slight differences in technology or costs can make all the difference in the world to company profitability. We've got our team of computer wizards who are on top of all that. Ironically, though, the same knowledge advantage that has allowed us to beat the bigger banks in this market makes them too scared to buy loans from us directly. They don't think that they can tell the good borrowers from the bad ones as well as we can. And they are afraid that we would pawn off our losers on them. The bottom line is that we would have to either sell the loans for a lot less than they are really worth, or else provide a lot of credit enhancement. (*See Exhibit D.*)

Langston: Well, I want the two of you to check into this in detail, particularly the possibility of securitizing some of our loan portfolio. Find out what our options are, and get back to me next week.

[*The meeting ends. Lana Tanner and Hugh Akston start working on the options available to the bank. Tanner will look into various kinds of securitization markets, while Akston will check into accounting and regulatory concerns.*]

The Second Meeting

Langston: Well, Lana, you were going to look into the possibility of more direct securitization. Do you think we could package several of our loans into a pool for sale as asset-backed bonds?

Tanner: Unfortunately, even though the companies we lend to produce very similar products for sale, they vary widely in financial structure, costs, and so forth. For example, Gell Microsystems has a long-term contract with Sintel to buy memory chips at present prices. So if Washington imposes tariffs on Japanese chips, thus raising the overall cost of memory chips, Gell will benefit. Other companies have a variety of long-term or short-term contracts with other parts of the industry, including some really complicated software licensing arrangements. Quite simply put, these companies are not very homogeneous. (*See Exhibit E.*) About the only thing they have in common is their industry and the fact that they do a significant portion of their sales through business lease agreements.

Langston: So what's the bottom line?

Tanner: Well, we could try to package some of our loans to these companies, but the very different payoffs between these firms would mean that investors would have to know a lot about which firms were in the pool. The Wall Street people I talked to seemed to think that we would lose a lot of basis points in trying to sell these things, or else that we would have to overcollateralize or buy insurance bonds.

Langston: Well, it seems that we are caught between a rock and a hard place. On the one hand, we have the ability to identify and monitor good loan situations better than anyone else can. But this very advantage makes it difficult for us to get these loans off of our books.

Akston: Also, there's a factor that we haven't talked about yet. As you mentioned, we are really heavily concentrated in one relatively narrow industry. Additionally, the personal computer industry has historically been very sensitive to macroeconomic changes. When the economy turns sour, individuals and businesses can easily postpone the purchase of a new computer. That means that all of the companies we fund probably have highly correlated patterns of default. Many of our individual customers work in the computer business, too, so that default patterns on our consumer loans are also correlated with the computer industry. If we securitize with recourse, we really haven't diversified ourselves against an overall change in the computer business climate. A downturn would only be temporary, and we've weathered them before, but if we don't get some real diversification, we will be much more leveraged with respect to this particular risk—and we might lose strategic market share to better-diversified banks during a tight squeeze.

Langston: Ouch! Well, this is tough. Hugh, tell me what you found out.

Akston: Well, one thing is clear. Unless we are willing to sell these loans at huge discounts, we will have to retain a significant amount of recourse. What that means is that, for RAP purposes especially, we will have a hard time getting this stuff off our books.

Langston: O.K. Well, here's what I'd like the two of you to do. We've discussed a lot of different options for the bank, but it is time to act. I'd like you to work together and systematically examine each of the options we've mentioned as well as any others you can think of. Report back to me what the pros and cons are of each approach, and give me your final recommendation about which policy would be best for us.

The Assignment

Present the pros and cons of each approach mentioned in the case, being sure to cover the issues of fundamental importance. Try to think creatively about alternative solutions to come up with something potentially better for the bank.

Exhibit A
Lone Star Bank's Business Loan Portfolio by Industry

	2000	2001	2002	2003	2004	2005
Construction	23%	28%	35%	38%	27%	23%
Computer Manufacturing	32%	38%	43%	475	52%	55%
Retail Stores	17%	18%	12%	10%	13%	12%
Distributors	5%	5%	4%	4%	5%	4%
Other	23%	11%	6%	1%	3%	6%

Exhibit B
Lone Star Bank
Profitability Measures

	2000	2001	2002	2003	2004	2005
Return on Assets	.900	.875	.000	.834	.752	.654
Return on Equity	11.2	11.0	0.0	9.4	8.7	6.32
Gross Margin	54.8	54.2	50.1	54.5	53.0	51.2

Exhibit C
Lone Star Bank
Construction/Real Estate Development Loans

	2000	2001	2002	2003	2004	2005
Volume (millions)	46.0	61.6	84.7	101.1	82.6	80.9
Loan Losses (percent)	.9	1.1	2.1	3.4	1.9	1.1
Estimated Net Return	0.02	0.01	−0.02	−0.02	0.00	0.01

Exhibit D
Estimated Loan Sale Values

Estimated values based on a static pass-through $20 million face value pool. Market values are based on an average of estimates obtained from investment bankers contacted.

True value (NPV of estimated future cash flows):	$17,200,000
Loan sale without recourse (no credit enhancements):	$16,100,000
Loan sale with full recourse:	$17,200,000
Loan sale without recourse (with credit enhancements with a cost of $700,000):	$17,100,000

Exhibit E
Customer Profiles

Company name	Sales	Employees (Millions of dollars)	D/A	ROA	P/E
Gell Microsystems	523	2100	.40	.21	23
Encore Systems	215	1450	.80	.19	*
Southgate Comp. Systems	207	934	.67	.23	31
Texlon	185	1200	.32	.13	18
ZEON	127	600	.87	.29	42

*Not publicly traded.
Note: D/A is debt/total assets in book value terms.
ROA is return on assets.
P/E is the ratio of stock price to reported earnings.

Review Questions

1. What are the similarities and the differences between loan sales and securitization?
2. What are the four basic components of a lending transaction? Why were these unified in the first place and why are they being decomposed now through securitization?
3. Discuss pass-through (both static and dynamic pool), asset-backed bonds, and pay-through, with particular focus on the differences between these contracts.

4. What are CMOs and REMICs?
5. Why are pass-throughs more popular than pay-throughs, and why are REMICs now replacing pass-throughs?
6. What are I/O - P/O strips and how can they be used to hedge interest-rate risk?
7. What are the supply- and demand-side forces stimulating securitization?
8. What is asset-backed commercial paper? Why has it become popular? Why don't corporations avoid banks and directly issue secured commercial paper?
9. Are there any "natural" limits to securitization? What are these? What sort of assets are most likely to be securitized and what assets are likely to be securitized?
10. Explain how a financial institution can use securitization as a tactical tool for balance sheet management and pricing, and as a strategic weapon for market penetration and diversification.
11. What does the anticipated future growth of securitization portend for the viability of banks and the ability of the Fed to control monetary aggregates?
12. Suppose bank A has two loans, each of which is due to be repaid one period hence and whose cash flows are independent and identically distributed random variables. Each loan will repay \$250 to the bank with probability 0.8 and \$125 with probability 0.2. However, while bank A knows this, prospective investors cannot distinguish this bank's loan portfolio from that of bank B that has the same number of loans, but each of its loans will repay \$250 with probability 0.5 and \$125 with probability 0.5. The prior belief of investors is that there is a 0.4 probability that bank A has the higher-valued portfolio and 0.6 probability that it has the lower-valued portfolio. Suppose that bank A wishes to securitize these loans, and it knows that if it does so without credit enhancement, the cost of communicating the true value of its loans to investors is 8 percent of the true value. Explore bank A's securitization alternatives. Assuming that a credit enhancer is available and that the credit enhancer could (at negligible cost) determine the true value of the loan portfolio, what sort of credit enhancement should bank A purchase? Assume everybody is risk neutral and that the discount rate is zero.
13. Given below is a slightly modified excerpt from "A Friendly Conversation." Critique it.

Moderator: O.K.! That's one for you, Alex. But I don't understand one thing. If banks are allowed to invest only in very safe assets, what happens to all of the assets that banks currently fund?

Appleton: No big deal. These can be shifted to the capital market or funded with uninsured deposits.

Moderator: But is such disintermediation or reintermediation necessarily a good thing?

Appleton: I don't see why not. Banks are already securitizing many of their assets, from credit-card receivables to mortgages. What I'm suggesting is only a natural extension of that process.

Butterworth: Sure, but there are natural limits to securitization. Besides, even with securitization, the bank acts as an originator. What you're proposing, Alex, is based, I think, on the premise that there is really nothing special about banks.

References

Anson, Weston, "Intangible Assets: A New Source of Security and Securitization," *The Secured Lender*, July–August 2005.

Ayotte, Kenneth, M., and Stav Gaon, "Costs and Benefits of 'Bankruptcy Remoteness,'" Working Paper, Columbia University Business School, May 2006.

Bagehot, Walter, *Lombard Street: A Description of the Money Market*, Kegan Paul, Trench, Trubner & Co., London, 1873.

Benveniste, Lawrence M., and Allen N. Berger, "Securitization with Recourse: An Instrument That Offers Uninsured Bank Depositors Sequential Claim," *Journal of Banking and Finance* 11–3, September 1987, 403–424.

Bevier, Michael, "Benefits of Taking a Longer-Term, Strategic View of Securitization," *Special Supplement on Asset Securitization, American Banker*, June 8, 1990.

Board of Governors of the Federal Reserve System, *Annual Statistical Digest*, 1988.

Boot, Arnoud W.A., and Anjan V. Thakor, "Security Design," *Journal of Finance*, 48–3, September 1993, 1349–1378.

Boyd, John H., and Bruce D. Smith, "Intermediation and the Equilibrium Allocation of Investment Capital: Implications for Economic Development," Working Paper, Federal Reserve Bank of Minneapolis, May 1993.

Caouette, John B., "As the Capital Markets Unbundle, What Will the Future Bring?" *Special Supplement on Asset Securitization, American Banker*, June 8, 1990.

Carney, Owen, and Ellen Starr, "Reporting and Regulatory Treatment of Asset Sales With Recourse," *Special Supplement on Asset Securitization, American Banker*, June 8, 1990.

Cocheo, Steve, "Securitization Option Blossom," *ABA Banking Journal*, August 1993, 45–48.

Cutler, Steward, and Paul Sveen, "Asset-Backed Commercial Paper," *Guide to International Commercial Paper*, January 1993, 5–12.

Dougherty, Ann J., "Orchestrating New Securities," *Secondary Mortgage Markets*, Spring 1987.

Federal Reserve Bulletin, February 1993.

Financial Accounting Standards Board, "Reporting by Transferors for Transfers of Receivables with Recourse," Statement of Financial Accounting Standards No. 77, December 1983.

Fishman, Michael, and Leon Kendall, *A Primer on Securitization*, The MIT Press, July 2000.

Gorton, Gary, and George Pennacchi, "Banks and Loan Sale: Marketing Nonmarketable Assets," Manuscript, University of Illinois – University of Pennsylvania, July 1993.

——, "Financial Intermediaries and Liquidity Creation," *Journal of Finance* 45, 1990, 49–72.

——, "Security Baskets and Index-Linked Securities," Working Paper, the Wharton School, University of Pennsylvania, May 1991.

Gorton, Gary, and Joseph G. Haubrich, "Loan Sales, Recourse, and Reputation: An Analysis of Secondary Loan Participations," Working Paper, the Wharton School, University of Pennsylvania, 1987.

Greenbaum, Stuart I., "Securitization, Asset Quality and Regulatory Reform," BRC Working Paper #147, Northwestern University, April 1987.

Greenbaum, Stuart I., and Anjan V. Thakor, "Bank Funding Modes: Securitization Versus Deposits," *Journal of Banking and Finance* 11–3, September 1987, 379–402.

Grossman, Sanford J., "An Introduction to the Theory of Rational Expectations Under Asymmetric Information," *Review of Economic Studies* 48, 1981, 541–559.

Haidorfer, Peter R., "Resolving the Conflict Regionalization Versus Diversification," *Special Supplement on Asset Securitization, American Banker*, June 8, 1990.

Huber, Stephen K., *Bank Officer's Handbook of Government Regulation*, Warren Gorham Lamont, 1992 Cumulative Supplement No. 2.

James, Christopher M., "The Use of Loan Sales and Standby Letters of Credit by Commercial Banks," *Journal of Monetary Economics* 22, 1988, 395–422.

Kareken, John H., "The Emergence and Regulation of Contingent Commitment Banking," *Journal of Banking and Finance* 11–3, September 1987, 359–378.

Kasper, Thomas A., "CMOs Fix Mismatches," *Savings Institutions*, January 1985.

Kraus, James R., "Capital Rules Spur Foreign Banks Into Asset-Backed Commercial Paper," *American Banker*, May 20, 1993.

Lelogeais, Jean Louis, "Asset Securitization: Strategic Initiative or Tactical Weapon?" *Special Supplement on Asset Securitization, American Banker*, June 8, 1990.

Martin, Daryl, and David C. Drews, "Intellectual Property: Collateral for Securitization or Lending, " *The Secured Lender*, July–August 2005.

Mester, Loretta J., "Traditional and Nontraditional Banking: An Information-Theoretic Approach," *Journal of Banking and Finance* 16–3, June 1992, 545–566.

Monahan, Maureen, "An Investor's Guide to Asset-Backed Securities," Monograph, Shearson Lehman Hutton, Inc., March 1989.

Pavel, Christine A., "Securitization," *Economic Perspectives*, Federal Reserve Bank of Chicago, July/August 1986, 16–31.

——., *Securitization*, Probus Publishing, Chicago, 1989.

Pavel, Christine, and David Phillis, "To Sell or Not to Sell: Loan Sales by Commercial Banks," Mimeo, Federal Reserve Bank of Chicago, 1987.

Rehm, Barbara A., "Fed Proposal Seen Spurring Asset Sales," *American Banker*, December 17, 1993.

Schwartz, Gilbert T., "Regulatory Issues in Connection With Securitization of Assets," *Special Supplement on Asset Securitization, American Banker*, June 8, 1990.

Shah, Salman, and Anjan V. Thakor, "Optimal Capital Structure and Project Financing," *Journal of Economic Theory* 42, August 1987, 209–243.

Shapiro, Harvey D., "The Securitization of Practically Everything," *Institutional Investor* 19, May 1985, 201.

Subrahmanyam, Avanidhar, "A Theory of Trading in Stock Index Futures," *Review of Financial Studies* 4–1, 1991, 17–52.

Thomas, Hugh, "The Effects of Asset Securitization on Seller Claimants," *Journal of Financial Intermediation* 10, 3–4, 2001, 306–330.

PART • V

The Deposit Contract

This page intentionally left blank

CHAPTER ♦ 10

The Deposit Contract and Insurance

"As to guaranteeing bank deposits, the minute the government starts to do that, . . . the government runs into a probable loss. We do not wish to make the United States government liable for the mistakes and errors of individual banks, and put a premium on unsound banking in the future."

Franklin Delano Roosevelt, in his first press conference as President of the United States

Glossary of Terms

Charter Value: The economic value of a bank to its owners (the shareholders). It can be viewed as the net present value of the profits expected to accrue to the shareholders over the life of the bank.

Null Hypothesis: In statistical-decision theory, when we believe something is true, we formulate the null hypothesis as the alternative to what we believe is true. Thus, when we perform statistical tests using the available data, we *expect* to reject the null hypothesis.

Anticompetitive Restrictions: Restrictions aimed at limiting competition in the banking industry.

Price Elasticity of Demand: A measure of the responsiveness of market demand to changes in price.

Junk Bonds: Very high (default) risk bonds issued by corporations. These bonds have low credit ratings and carry high yields.

Capital Asset Pricing Model: A model describing how risk is priced in the capital market. In particular, it predicts a linear relationship between the expected return on a security and its systematic risk factor (defined as "beta," the ratio of the covariance of the return on the security with that of the market to the variance of the market return).

Introduction

In earlier chapters, we focused on the asset side of the balance sheet for depository institutions. We now shift to the liability side. Although depository institutions have a wide variety of liabilities, in this chapter we will concentrate on different types of deposits and we will turn to capital in Chapter 11.

In the United States, the terms "bank deposits" and "deposit insurance" are almost inseparable. Yet, it is essential to distinguish the issues raised by the deposit contract *per se* from those related to deposit insurance. Thus, we will first discuss the deposit contract without the insurance aspect. We will then discuss "liability management," which is the process of managing the bank's net interest margin, that is the difference between the asset revenues and the liability costs, expressed as a fraction of total assets. This will be followed by an analysis of deposit insurance. Having previously discussed the uninsured deposit contract, we will be able to see how governmental deposit insurance alters the deposit contract, and the behaviors of deposit takers. This, in turn, sets the stage for an analysis of reform proposals in the next chapter.

We doubt that anyone remains to be convinced about the importance of deposit insurance-related issues. There is an almost surreal air about the scandalous 1980s. According to the 1993 *Economic Report of the President*, the S&L industry lost between $100 billion to $160 billion. Commercial banking was shaken to its foundations. Fundamental regulatory reform followed, and a transformation of the financial services industry has occurred as a result.

Many have blamed deposit insurance and greed for the S&L crisis and the widespread banking failures of the 1980s. While this seems to be accepted, it is more difficult to explain why we have deposit insurance, and in particular, why we have the kind of deposit contract that seems to make federal insurance desirable. Discussions of these issues figure prominently in this chapter. Recent events have taught us many valuable lessons. What is unfortunate is that the ongoing crisis was, to a great extent, avoidable, and the regulatory reforms that followed the crisis made sense well before the crisis occurred. As early as 1977, academic publications made the point that federally insured depository institutions had powerful incentives to take asset risk that was excessive from a social welfare standpoint, and that capital regulation, as it existed then, by itself was incapable of controlling these incentives, so that a fundamental reform of regulation was necessary.[1,2]

Some might argue, however, that our historical experience (particularly since the advent of federal deposit insurance, following the Great Depression) did not prepare us for the systemic shocks of the last decade. In the post-1933 period, extremely low bank failure rates[3] made banking a rather unusual industry. So another puzzle is: Why the rash of failures did not occur prior to the 1980s? It turns out that the empirical and theoretical research on which this chapter is based provides valuable insights into the timing of the recent difficulties, and leads us to conclude that, despite our comfortable post-Depression experience, we should have foreseen many of the things that happened.

1. See Kahane (1977) and Merton (1977).
2. See *Capital Issues in Banking* published by the *Association of Reserve City Bankers* (1988).
3. These failure rates were less than 0.3 percent.

The rest of this chapter is organized as follows. In the next section, we discuss the deposit contract. After that, we take up liability management and how it has been affected by interest-rate deregulation and deposit insurance. Then we discuss deposit insurance. We examine the arguments for and against deposit insurance, including the ability of governmentally provided deposit insurance to ward off runs on banks and panics. Issues related to the risk-sensitive pricing of deposit insurance are also examined, as is an analysis of the empirical evidence on the importance of moral hazard in federally insured depository institutions. The empirical evidence also provides insights into the timing of problems with deposit insurance. We then discuss the 1980s deposit insurance debacle in the United States, and developments that have occurred since then.

The Deposit Contract

The Nature of the Deposit Contract

Deposit contracts either have *defined maturities* like certificates of deposit (CDs), or are *withdrawable on demand*. We will focus on demand deposits, the quintessential banking liability. A demand deposit is created when an individual or firm deposits money in an account from which this money can be withdrawn at a moment's notice, that is, on demand.

The demand deposit contract has four important features:

- Its maturity is infinitesimal and can be rolled over indefinitely.
- It is a debt contract.
- It is not traded in a secondary market.
- It is governed by a "sequential service" constraint.

Maturity: The maturity is such that the depositor is promised the ability to withdraw at any time without penalty, that is, the depositor can sell the bank's liability back to the bank at par. Thus, a demand deposit is virtually as liquid as currency. The key difference is that currency carries no default risk, whereas an uninsured bank could default and not be able to fully satisfy withdrawal demands. Indeed, throughout *this* section we will assume that there is no deposit insurance, so that we can focus on the characteristics of the deposit contract itself.

Debt Contract: Because the deposit is a debt contract, the depositor in an uninsured bank confronts the same asset-substitution moral hazard in dealing with the bank as the bank does in dealing with its borrowers (recall Chapters 5 and 6). That is, when a bank creates a deposit, it is simply borrowing from the depositor.

Nontraded Contract: The fact that demand deposits are not traded in a secondary market implies that the depositor's payoff does not depend directly on how information about the bank is processed by other market participants, that is, the depositor does not face market price risk. Unlike a person who plans to sell a traded security in the market at the (random) price prevailing at a future date, a demand depositor knows precisely (in nominal terms) how much she will receive at *any* future point in time when she withdraws from her account, subject to the condition that the bank is solvent.

This last condition is not always satisfied, however. In fact, if things were believed to be going badly for the bank, we would expect the suspicious depositors to rush to the bank to withdraw their deposits. If you arrive late, it is possible that in paying off the earlier depositors the bank will have run out of money by the time you get there. In this case, absent deposit insurance, the maximum amount you can withdraw would be less than you had anticipated. In this sense, your payoff depends on what other depositors believe about the bank, just as it does with any traded debt contract that you liquidate prior to maturity.

The Sequential Service Constraint: This dependence of your payoff on the actions of other depositors occurs because the deposit contract satisfies a sequential service constraint (SSC). Hence, when a depositor seeks to withdraw, the amount the bank pays depends only on what was promised and on his place in the queue of depositors wishing to withdraw. In particular, the depositor's payoff cannot depend on any information that the bank may have about depositors in the queue behind that depositor. Thus, the bank pays depositors on a "first come, first served" basis. To see this, consider a bank that has \$5 in equity, and \$95 in interest-free deposits acquired from 95 depositors (each of whom deposited \$1). The bank's \$100 of assets consist of \$20 in cash and loans that are currently worth \$80 if held to maturity. But if the loans are prematurely liquidated, they are worth only \$27.50. Thus, the current (premature) liquidation value of the bank is \$47.50. Now imagine that some depositors rush to withdraw their money. Others hear about this and become suspicious about the bank's assets. There is now a full-scale bank run. You are the 48th depositor in a queue of 95 when the bank's doors open in the morning. As the branch manager walks in, she counts the number of people in the queue and sees that every depositor is there to withdraw. Despite this, the SSC dictates that the bank cannot use this information in determining how much the first-in-line depositor should be paid. In this case, the manager is forced to call the outstanding loans, that is, liquidate them to collect \$27.50. The first 47 depositors will each receive 41. You will receive \$0.50, and all those behind you will go home empty-handed. One might argue that a more equitable approach would have been to give each of the 95 waiting depositors \$0.50. But the SSC precludes that.

The nature of the deposit contract is worth examining for two reasons. First, when *all* of the bank's liabilities are uninsured, these features have significant implications for the disciplining of bank management. This suggests that the details of the demand deposit contract are probably not an outcome of chance; they serve a purpose. Second, when deposits are insured, some of these features of the demand deposit contract *encourage* bank runs, thus increasing the liability of the deposit insurer.

The Demand Deposit Contract and Economic Incentives

The Effects of Nontradability and the Debt-Like Nature of Deposits: Consider first that demandable debt is not traded and that it is a debt contract. The analysis in Chapters 5 and 6 implies that the depository institution in this case has an incentive to increase asset risk to the detriment of the depositors. That is, the institution's managers have an incentive to invest in risky loans that transfer wealth from depositors to shareholders. Similarly, depositors face moral hazard in that the institution has an incentive to shirk in monitoring the borrowers to whom it has extended loans. This too adversely affects the depositors' expected payoff. A third form of moral hazard is fraud. Deposits are essentially "someone else's" money, and managers may be tempted to appropriate some of that money for themselves. While these pathologies have been attributed to federal deposit insurance, they were encountered even prior to the adoption of deposit insurance,[4] and our theory predicts that incentives for managerial fraud exist even with (nontraded) deposits that are uninsured. That the deposit contract is not traded aggravates the moral hazard problem because the discipline imposed by market pricing is absent.

The Effect of Maturity: It turns out, however, that the other two features of the demand deposit contract—its infinitesimal maturity and the SSC—help to attenuate these different types of moral hazard. In developing the intuition below,[5] we first consider the effect of the undefined maturity.

Suppose that there are numerous individuals who demand deposit accounts at a bank. It is natural to expect that some of these depositors are particularly skilled in analyzing the bank's financial health, whereas others are less able. Let us suppose that these skilled depositors keep a watchful eye on the bank's managers because they recognize that moral hazard could diminish their expected payoff. Now, imagine that a few of these vigilant depositors discover that the bank's risky loans are not doing well. Default on many of these loans is likely. Moreover, these depositors discover that the bank has extended numerous loans to close friends of the top managers; this raises suspicion of fraud. What should these informed depositors do? Since they have information that the bank is in peril and may default on its deposit obligations, their best bet is to withdraw their funds as quickly as possible.

When these informed depositors withdraw their funds from the bank, there are two possibilities. One is that the uninformed depositors do not react. In this case, the total outflow of funds from the bank will depend on the size of the deposit holdings of the informed depositors. If their holdings are large enough, the bank will be compelled to attract new deposits. The second possibility is that some or all of the uninformed depositors observe the withdrawals of the informed depositors and decide to follow suit. In this case, there is a bank run. In either case, the bank will need to attract new deposits to replace withdrawals, or liquidate. Liquidation will involve either the calling back of loans, with the associated disruptions in the productive activities of borrowers, or loan sales to other banks. The alternative of attracting new deposits will be difficult, for obvious reasons. Prospective depositors will see the large deposit withdrawals and will be reluctant to entrust their money to the bank. And even if some deposit money flows in, the bank will need to pay higher interest rates on these deposits. Thus, deposit withdrawals by the informed depositors are likely to be costly to the bank. The anticipation of incurring these costs could deter the bank's managers from risky investments, and from shirking on the monitoring of borrowers. It could also reduce the temptation to defraud the depositors.

The Role of the SSC: This argument suggests that the demandable nature of deposits helps to keep bank management on its toes. There is a slight hitch in this disciplining process, however. If a depositor can rely on other depositors to monitor

4. See Calomiris and Kahn (1991).

5. This intuition is based on Calomiris and Kahn (1991), and Calomiris, Kahn, and Krasa (1991). See also Diamond and Rajan (2001), Jacklin (1987, 1989) and Jacklin and Bhattacharya (1988).

the bank, then all that such a depositor has to do is to keep an eye on the informed depositors. There is no need for the "free-riding" depositor to expend personal resources to monitor the bank. This can subvert depositor monitoring. The reason is that *every* depositor may think that others will do the necessary monitoring, and in that case, no one monitors! This is where the SSC comes into play. Because a depositor's expected payoff is greater if he is at the front of the queue than if he is at the rear, he recognizes that by playing a "follow the leader" strategy, his expected payoff is lower than if he monitors himself. This strengthens each individual depositor's incentive to monitor. These ideas are made concrete in the example developed in the box below.

An Illustration of the Incentive Effects of the [Uninsured] Deposit Contract

Example 10.1 Consider a bank that receives a \$1 deposit at $t = 0$ from each of 105 different depositors. It invests \$10 of shareholders' equity in the bank and lends \$110, keeping \$5 as cash reserves. Out of the 105 depositors, there are 30 depositors (called type-D_1 depositors) who are capable of monitoring the bank's management; the remaining depositors (called type-D_2 depositors) keep their money in the bank simply for transactions and safekeeping purposes. The cost of monitoring the bank for an individual type-D_1 depositor is \$0.01 per period.

The bank has two mutually exclusive investment opportunities. Project (or loan) A pays \$200 with probability 0.7 and zero with probability 0.3 at $t = 1$. Project B pays \$150 with probability 0.9 and \$112 with probability 0.1 at $t = 1$. If the bank chooses one of these projects, the probability that the bank will actually end up with that project is 0.9. With probability 0.1, the bank will have inadvertently chosen the other project. Thus, we assume that the bank may make errors in project choice.[1] By monitoring the bank, a type-D_1 depositor can discover the bank's true project choice at some point in time intermediate between $t = 0$ and $t = 1$, say at $t = 1/2$. These depositors can, if they desire, force liquidation of the bank by withdrawing their deposits at $t = 1/2$, and the threat of this liquidation provides a disincentive to the bank to choose the risky project. Note that the bank's projects or loans mature at $t = 1$. If they are liquidated at $t = 1/2$, they are worth only \$25 to the bank. Under the terms of the deposit contract, the bank promises to pay a 12 percent interest (conditional on the bank having the financial capacity to do so) if deposit withdrawal occurs at $t = 1$, and no interest if withdrawal occurs before that. Thus, a depositor is entitled to \$1.12 if she withdraws at $t = 1$, and \$1 if she withdraws at $t = 1/2$. The risk-free discount rate is zero and all agents are risk neutral.

All the type-D_2 depositors plan to withdraw at $t = 1$, but each is subject to a random liquidity-motivated desire to withdraw at $t = 1/2$. To simplify, we will assume that even though no one knows in advance which (type-D_2) depositors will wish to withdraw at $t = 1/2$, the *fraction* of those who will wish to withdraw is known to be 5/75. That is, five type-D_2 depositors will wish to withdraw at $t = 1/2$. Assume that the bank's managers make decisions in the best interests of their shareholders. Compute the equilibrium strategies of the bank and its depositors.

Solution It is useful to summarize the strategies available to the bank and the different types of depositors before we begin to analyze the solution. These are listed below.

TABLE 10.1 Strategies of Participants

Agent	Strategies		
Bank	Choose project A at t = 0	Choose project B at t = 0	
Type-D_1 depositors	Monitor and decide whether or not to withdraw at t = 1/2 based upon result of monitoring	Do not monitor and withdraw at t = 1	Do not monitor and withdraw at t = 1/2;
Liquidity-motivated Type-D_2 depositors	Withdraw at t = 1/2	Withdraw at t = 1/2	
Other (Patient) Type-D_2 depositors	Withdraw at t = 1	Withdraw at t = 1	

We will solve this problem in four steps. First, we analyze the bank's project choice in the case in which the type-D_1 depositors do not monitor and the bank knows that there is no monitoring. We show that the bank chooses project A in this case. Second, we show that our assumption in Step 1 is invalid because it cannot be a Nash equilibrium for no type-D_1 depositors to monitor. Next, we wish to examine if it is a Nash equilibrium for all the type-D_1 depositors to monitor. We do this in two steps. In Step 3, we show that the bank chooses project B if it believes that all the type-D_1 depositors will monitor. Then in Step 4, we examine the strategy of a type-D_1 depositor when he knows that all the other type-D_1 depositors will monitor and the bank has opted for project B. We show that this type-D_1 depositor *will* wish to monitor. This verifies that it is indeed a Nash equilibrium for all the type-D_1 depositors to monitor.

The key assumption in this example is that the bank's project choice cannot be contracted upon because not all depositors can observe it. If this were not the case, there would be no role for depositor monitoring.

Step 1 We will first analyze the outcome in which the type-D_1 depositors do *not* monitor the bank. Given that the bank knows that there is no monitoring, which project will it prefer? If it chooses project A and if this choice is error-free, the expected payoff of its shareholders is

$$\underset{\substack{\uparrow\\ \text{probability}\\ \text{of success}}}{0.7} \times \underset{\substack{\uparrow\\ \text{total}\\ \text{payoff}}}{[200} - \underset{\substack{\uparrow\\ \text{bank's repayment}\\ \text{to its depositors}}}{112]} = \$61.6,$$

and if it chooses project B and this choice is error-free, the expected payoff is

(*Continued*)

$$\underset{\substack{\uparrow\\ \text{probability}\\ \text{of high}\\ \text{payoff}}}{0.9} \times \underset{\substack{\uparrow\\ \text{high payoff}\\ \text{on}\\ \text{project B}}}{[150} - \underset{\substack{\uparrow\\ \text{bank's}\\ \text{repayment}\\ \text{to depositors}}}{112]} + \underset{\substack{\uparrow\\ \text{probability}\\ \text{of low}\\ \text{payoff}}}{0.1} \times \underset{\substack{\uparrow\\ \text{low payoff}\\ \text{on}\\ \text{project B}}}{[112} - \underset{\substack{\uparrow\\ \text{repayment}\\ \text{to depositors}}}{112]} = \$3$$

When project choice is error-prone, the expected payoff of the bank's shareholders when project A is chosen is

$$0.9 \times \$61.6 + 0.1 \times \$34.2 = \$58.86,$$

and their expected payoff when project B is chosen is

$$0.9 \times \$34.2 + 0.1 \times \$61.6 = \$36.94.$$

Thus, if there is no monitoring, the bank will choose project A.

Step 2 The question now is: Can it be a Nash equilibrium for *no* type-D_1 depositor to monitor? This is equivalent to asking whether it is in the best interest of every individual type-D_1 depositor not to monitor when she knows that no other type-D_1 depositors are monitoring. Suppose you are one of those type-D_1 depositors. If you do not monitor, your expected payoff is

$$\underset{\substack{\uparrow\\ \text{probability}\\ \text{that bank}\\ \text{will actually}\\ \text{have project A}}}{0.9} \times \underset{\substack{\uparrow\\ \text{depositor's}\\ \text{expected payoff}\\ \text{if bank chooses}\\ \text{project A}}}{[0.7 \times \$1.12]} + \underset{\substack{\uparrow\\ \text{probability}\\ \text{that bank}\\ \text{inadvertently}\\ \text{chose project B}}}{0.1} \times \underset{\substack{\uparrow\\ \text{depositor's}\\ \text{payoff when}\\ \text{project B is chosen}}}{[\$1.12]} = \$0.8176.$$

Now, if you do monitor, and discover that the bank chose project A, what should you do? If you do nothing (that is, you do not withdraw your deposit), your expected payoff at $t = 1$ is

$$0.7 \times \$1.12 - \underset{\substack{\uparrow\\ \text{your monitoring cost}}}{\$0.01} = \$0.774.$$

If you withdraw, you know that the bank will be forced to liquidate its asset portfolio since it has only \$5 in cash reserves and there are five type-D_2 depositors who will withdraw at $t = 1/2$ for liquidity purposes. Liquidation will fetch \$25, so that the bank will have a total of \$30 to disburse. You are sure to receive your \$1 at $t = 1/2$. Thus, your payoff will be

$$\$1 - \$0.01 = \$0.99.$$

This means that if you monitor and discover that project A has been chosen, you should demand to withdraw your deposit at $t = 1/2$. On the other hand, if you find that project B was chosen, your payoff is

$$\$1.12 - \$0.01 = \$1.11$$

if you wait until $t = 1$ to withdraw, and it is \$0.99 if you withdraw at $t = 1/2$. Hence, it is better for you to wait until $t = 1$ (remember that your *time value* of money between $t = 1/2$ and $t = 1$ is zero). We can now compute the overall expected payoff to you from monitoring. This payoff is

$$\underset{\substack{\uparrow \\ \text{probability} \\ \text{that you will} \\ \text{discover project A} \\ \text{was chosen and will} \\ \text{therefore withdraw} \\ \text{at } t = 1/2}}{0.9} \times \$0.99 + \underset{\substack{\uparrow \\ \text{probability that you} \\ \text{will discover} \\ \text{that project B was chosen} \\ \text{and will therefore withdraw} \\ \text{at } t = 1}}{0.1} \times \$1.11 = \$1.002.$$

Clearly, this payoff exceeds your payoff if you do not monitor (\$0.8176). This proves that you have an incentive to monitor when others do not, which means that it cannot be a Nash equilibrium for nobody to monitor.

Step 3 Let us now examine if it is a Nash equilibrium for all the type-D_1 depositors to monitor. We begin by noting that if the bank believes that all these depositors will monitor, then it is in the bank's best interest to choose project B. This is verified below.

If the bank chooses project A, then there is only a 0.1 probability that project B will be inadvertently chosen. That is, there is a 0.9 probability that the bank will be liquidated at $t = 1/2$. Thus, the expected payoff of the bank's shareholders from opting for project A is

$$\underset{\substack{\uparrow \\ \text{probability that the} \\ \text{bank will not be} \\ \text{liquidated}}}{0.1} \times \underset{\substack{\uparrow \\ \text{expected payoff of bank's shareholders} \\ \text{if project B is (inadvertently) chosen} \\ \text{and bank is not liquidated}}}{\$34.2} = \$3.42.$$

If the bank opts for project B, then there is only a 0.1 probability that the bank will be liquidated at $t = 1/2$ (this is the probability that project A will be erroneously picked). Thus, the expected payoff of the bank's shareholders from opting for project B will be

$$0.9 \times \$34.2 = \$30.78.$$

Clearly, the shareholders are better off opting for project B.

(*Continued*)

Step 4 The next step is to examine the strategy of a type-D_1 depositor when he knows that all the other type-D_1 depositors will monitor and the bank has opted for project B. If you are that depositor and you monitor, your payoff (at $t = 1$) is \$1.12 if you discover at $t = 1/2$ that the bank indeed chose project B. But if you discover that project A was chosen, then you will want to withdraw your deposit. The problem now is a little different from the previous case. You realize that if you discover that project A was chosen, so will the 29 other type-D_1 depositors. When added to the five liquidity-motivated type-D_2 depositors, this means that the line of those who wish to withdraw at $t = 1/2$ will be 35 depositors long. But the bank has only \$30 upon liquidation, and hence can only satisfy the first 30 depositors. Assuming that each person who goes to the bank will have an equal probability of being one of the first 30, the probability is 30/35 that you will be one of the first 30 withdrawers.[2] In this case, your *expected* payoff is only $\frac{30}{35} \times \$1 = \$0.8571$, since you get nothing if you are not one of the first 30 in line. Thus, your overall expected payoff from monitoring is given by

$$0.9 \times \$1.12 + 0.1 \times \$0.8571 - \underbrace{\$0.01}_{\text{monitoring cost}} = \$1.0837.$$

If you decide not to monitor, then you are behaving like a type-D_2 depositor. Your expected payoff will be \$1.12 if the other type-D_1 depositors discover that project B was chosen (the probability of this is 0.9), and it will be zero if they discover that project A was chosen and decide to liquidate the bank at $t = 1/2$ (the probability of this is 0.1). Hence, your overall expected payoff from not monitoring is

$$0.9 \times \$1.12 = \$1.008.$$

Another possible strategy is for you to behave like a liquidity-motivated type-D_2 depositor and withdraw your deposit at $t = 1/2$ *without* monitoring. In this case, you recognize that there is a 0.9 probability that the other type-D_1 depositors will not withdraw and a 0.1 probability that they will. If the other type-D_1 depositors do not withdraw, there are only six depositors in all (including you) who wish to withdraw at $t = 1/2$. The bank will be forced to liquidate, and you will receive your \$1 for sure. If the other type-D_1 depositors withdraw, the bank will also liquidate, and you will have a 30/35 chance of getting your \$1. Thus, your expected payoff from withdrawing without monitoring is

$$0.9 \times 1 \times \$1 + 0.1 \times \frac{30}{35} \times \$1 = \$0.9086.$$

Comparing the three payoffs (\$1.0837, \$1.008, and \$0.9086), we see that your best strategy is to monitor. Hence, it is a Nash equilibrium for all the type-D_1 depositors to monitor the bank, and for the bank to choose project B.

1. This feature ensures that the type-D_1 depositors do monitor the bank in equilibrium. The reason is that the threat of depositor monitoring will, in equilibrium, cause the bank to choose the project desired by the depositors. If this choice were error-free, depositors would anticipate that the bank will make the desired project choice and therefore perceive no need to monitor. But then the bank, in turn, should anticipate the behavior of the depositors and decide to invest in the project preferred by its own shareholders. And so on and on! The point is that we have a time consistency problem that leads to there being no equilibrium. However, as our solution will make clear this problem can be avoided when the bank's project choice is error-prone.

2. By the SSC, this is the probability that you will receive your \$1.

Although we worked out this numerical example explicitly for the case of asset-substitution moral hazard, the intuition for managerial fraud is similar. In either case, the demandable nature of deposits puts pressure on bank management to not deviate too far from the desires of the depositors, and the SSC lends credibility to the depositors' threat to monitor to ensure "proper" bank behavior by creating a situation in which all vigilant depositors wish to monitor. Thus, these specific features of the deposit contract play an important role in aligning the incentives of the contracting parties in an uninsured bank. This leads naturally to the question of deposit insurance. Before we get to that, however, we discuss liability management in a bank.

Liability Management

We have thus far discussed the economics of the deposit contract. The use of the deposit contract is an integral component of what is called *liability management*.

What Is Liability Management?

Depository institutions pay particular attention to their *net-interest margin* (NIM), which is the difference between the yield on assets and the interest cost of liabilities, expressed as a fraction of total assets. Liability management refers to the institution's strategies for maintaining the continuity and cost effectiveness of funding assets.

There are three main (interrelated) issues in liability management. The first is *diversification*, which refers to choosing among funding sources so as to avoid over-dependence on a particular source. A second choice involves the *mix* of liabilities. Depository institutions raise funds using a variety of deposits, each of which represents a specific contractual form that is a strategic choice. The third choice is about liability *maturity structure*, which determines the bank's interest-rate risk exposure for a given asset maturity structure. We discuss each issue briefly in what follows.

Diversification

Diversifying funding sources reduces liquidity risk (recall Chapter 4). Borrowing and lending in the federal funds market, borrowing at the discount window, dealing with repurchase agreements, and utilizing large CDs, brokered deposits, and Eurodollar deposits are techniques that banks use to diversify. Borrowing in the fed funds market and at the discount window is usually short term; most fed funds transactions are *overnight loans*, although the number of *term fed funds* transactions, with maturities in weeks, has increased recently. For longer-maturity liabilities, banks rely on a variety of deposits. Prominent among these are *negotiable CDs*, called *jumbos*, which are actively traded large-denomination time deposits with market-determined interest rates, a minimum maturity of one week, and denominations exceeding $100,000. Most negotiable CDs are issued directly to customers, although some large institutions issue them to brokers, who then sell them to other investors. Deposits marketed this way are called *brokered deposits*.

Large banks also use *Eurodollar deposits*, which are time deposits denominated in dollars but held in banks outside the United States, including foreign branches of U.S. banks. Eurodeposits are created in many ways. Perhaps the simplest way is when an American transfers money on deposit in a U.S. bank to a bank in another country. These deposits remain in dollars. Eurodeposits are subject to the Federal Reserve's cash-asset reserve requirements, and are not protected by U.S. deposit insurance.

Banks also raise funds by using *repurchase agreements* or "repos." A repo is the sale of a marketable security, with the agreement to repurchase it at a specified future date, that is, it is a loan secured by a marketable security. As long as the securities pledged against repos are U.S. government or government agency securities, repos are not subject to reserve requirements. Repos range in maturity from overnight to a month or more. Since repos involve collateral, they are not considered deposits and hence are not covered by deposit insurance.

Banks use a variety of other funding sources, such as subordinated debt as well as securitization and loan sales. Securitization also facilitates diversification of the bank's loan portfolio. Moreover, bank holding companies can issue commercial paper.

Liability Mix

Bank liabilities can be divided into two categories: products and investment instruments.[6] A product entitles the purchaser to a financial claim as well as to some bank services. That is, it is a contract that bundles monetary and possibly nonmonetary payoffs. An example is a checking account on which the bank pays interest and provides transactions services. For corporations, other services include cash management at possibly subsidized prices. Thus, purchasers of product-based deposits, called "customers," receive both explicit and implicit interest, and the demand for such deposits depends both on the explicit interest as well as on the value depositors attach to the bank's services.[7] Because many of these services are demanded by retail depositors, deposits tend to be small (below the *de jure* deposit insurance coverage limit of $100,000 per account). Moreover, customers prefer to have the payoffs on their contracts as insensitive as possible to the fortunes of the intermediary itself. For example, a life insurance policy provides its beneficiaries with a specified cash payment conditional on the death of the insured. That function is less efficiently performed if the contract calls instead for the death benefit to be conditioned on the financial condition of the insurance company as well as on the death of the insured.[8] Consequently, an increase in the policyholder's risk due to a decline in the insurance company's financial condition may require a greater reduction in the insurance premium than would be actuarially fair. It may, therefore, pay for the insurance company to reduce the policyholder's risk as much as possible. In the case of banks, this may explain why product-based deposits are typically fully insured.

6. See Merton (1993).

7. These services are often valued very highly by depositors. Recent empirical evidence has shown that banks enjoy significant economic rents from money-market deposit accounts and NOW (Negotiable Orders of Withdrawal) accounts, both of which are retail deposit accounts. See Hutchison and Pennacchi (1992).

8. Merton (1993) suggests that an Arrow-Debreu economy (see the discussion of market incompleteness in Chapter 1) illustrates this point. A complete set of such securities provides a Pareto-efficient allocation of resources. But efficiency would be lost if the payoffs on such securities were also contingent on the issuer's financial condition [see also Merton (1989)].

Investment instruments, on the other hand, are simply financial claims, similar to the liabilities of nonfinancial firms. The bank provides no transactions or other services to the claimholder, so the design of these contracts involves the same risk-return trade-offs faced by nonfinancial firms. An example of an investment instrument is a brokered CD. Deposit contracts that are investment instruments tend to be purchased by institutions, are relatively large in denominations, and include uninsured deposits. Their prices are usually determined through secondary-market trading.

One of the bank's liability-management choices is the appropriate mix of product-based deposits and investment instruments. Because of the relative insensitivity of their values to the bank's riskiness, product-based deposits do not involve much monitoring of bank management by depositors. Investment instruments, on the other hand, have values that are sensitive to the bank's riskiness, and it pays for the holders of these claims to monitor the bank. The bank is, therefore, subject to greater market discipline with these deposits. From the standpoint of the bank's management, there may be a desire to reduce the bank's reliance on such deposits in order to limit market discipline. Of course, doing so may sacrifice diversification in funding sources, with the attendant liquidity risk that may eventually result in a loss of control for management. The bank's shareholders, on the other hand, would like sufficient reliance on investment instruments to ensure the desired level of market discipline. This suggests a liability-management agency problem between shareholders and managers of banks.

The Duration Structure

Given its asset duration structure, the bank's choice of liability duration structure will determine its interest-rate risk. Given long-duration assets, the bank faces a trade-off in making this choice. On the one hand, choosing a matching long duration on the liability side will minimize interest-rate risk. On the other hand, given the possibility of new information arrival, it may be efficient to choose a shorter duration structure and allow for periodic repricing of deposits. This can reduce the distortions that can arise from private information possessed by the bank at a particular time that may be released to the market later.[9]

Banks often resolve this tension by using derivatives (Chapter 8). The better-managed banks purchase the least expensive assets and liabilities and then use options, futures, and swaps to achieve the desired degree of immunization against interest-rate risk.

Deposit Insurance

The Rationale for Deposit Insurance: A Historical Perspective

The Need for Deposit Insurance: If the demand deposit contract discussed earlier works well in disciplining bank management, why do we need deposit insurance? The reasons are many. Not all make perfect sense in today's environment, but we will get to that later. For now, let us simply note that an uninsured (demand) deposit contract

9. See Flannery (1992).

can be quite disruptive. In a sense, it can lead to *overdisciplining* of banks. This can be seen as follows. In the previous section, we assumed for simplicity that the vigilant depositors could discover the bank's project choice without error. In reality, this discovery is likely to be error-prone. It is then possible that the bank is forced to liquidate assets even when its project choice is congruent with the preferences of depositors. This is socially wasteful *ex post*.

In addition, systematic elements in the risk profiles of the asset portfolios of banks may give risk to a *contagion effect* among banks. That is, when one bank fails, depositors suspect that the failure may be due to systematic risk elements that pervade the asset portfolios of *all* banks in that geographical area, and this may lead to spreading bank runs. Since it often takes a long time for the precise reasons for a bank's failure to become public, the contagion effect may be encountered even when the failure of a particular bank is due to idiosyncratic factors such as poor management. Indeed, this is the rationale for the "too big to fail" doctrine, which leads the government to rescue sufficiently large banks from failure.

Both of these problems are reduced with deposit insurance. When a government agency insures a bank's deposits, it guarantees that the depositors will receive their promised payment, regardless of the bank's financial condition. This makes it unnecessary for depositors to monitor the bank, and it lessens the likelihood of runs on individual banks or on groups of banks.

Historical Background: Federal deposit insurance came into existence in the United States with the enactment of the Banking Act of 1933, and the creation of the Federal Deposit Insurance Corporation (FDIC) to insure bank deposits. The insurance system was extended the following year to S&Ls with the creation of the Federal Savings and Loan Insurance Corporation (FSLIC), which insured S&L shares (deposits). In 1971, deposit insurance was also made available to credit unions.[10] All of this was inspired by the Great Depression and the massive runs on banks that forced President Roosevelt to declare a "banking holiday" in March of 1933. The banking panics of the Great Depression were not new, however. There were as many as seven panics from 1866 to 1934. We will use the term "bank run" (in the singular) to denote a situation in which depositors at a *single* bank wish to exchange their deposits for currency, and the term "banking panic" to "denote a situation in which depositors at many banks wish to exchange their deposits for currency.

Before federal deposit insurance, panics were often addressed by *suspending convertibility* of deposits into cash. Under this approach, the bank was simply closed to depositors who wished to withdraw their money. By giving the banks "breathing room" during which "mass hysteria" had a chance to die down, more information about the financial condition of the bank could be released. Unless this information confirmed the worst fears of depositors, they could be persuaded to refrain from withdrawing their money when the suspension was lifted. Suspension amounted to default on the deposit contract and was a violation of banking law. Nevertheless, five out of the seven panics referred to previously involved suspension of convertibility (those in 1873, 1890, 1893, 1907, and 1914).[11]

10. Legally, a credit union does not accept deposits but issues shares in the credit union to its members. In reality, credit union shares are so similar to deposits that we will not distinguish between them.

11. See Gorton (1988).

Another method that was used during banking panics was the *issuance of clearinghouse loan certificates*. These arose from *Commercial-Bank Clearinghouses* (CBCHs), private-market arrangements among banks that served some of the functions of a central bank. Initially a CBCH was formed to facilitate check clearing.[12] Prior to the formation of the New York CBCH in 1853, for example, commercial banks collected checks by a process of daily exchange and settlement with each other. The clearinghouse centralized the settlement process by permitting exchange to be made with the clearinghouse alone. However, as it evolved, the clearinghouse was able to provide additional information-based services such as *certification* (based on a minimum capital requirement needed to become a member of the clearinghouse) and *monitoring* (based on periodic audits) of its member banks. Members who failed to satisfy CBCH regulations were disciplined with fines or expulsions. This economized on individual monitoring costs that depositors would have had to incur in the absence of a clearinghouse.

One way for a bank to reduce the likelihood of a run is to reduce the depositors' concern about the bank's assets. The clearinghouse loan certificate, first issued during the panic of 1857, was an attempt to do this. A policy committee of the CBCH first authorized the issuance of loan certificates. Whenever a member bank had insufficient cash to satisfy deposit withdrawals, it could apply to the CBCH loan committee for certificates. Borrowing banks were charged interest rates varying from 6 to 7 percent and were required to present acceptable collateral. These certificates, which typically had maturities of 1 to 3 months, could be used by the bank in place of currency. Depositors were willing to accept the loan certificates in exchange for demand deposits because the loan certificates were claims on the CBCH, rather than on the individual bank. Thus, depositors obtained some insurance (diversification benefits) against individual bank failure. This meant that when there was a run on a bank, the bank could either pay off depositors in loan certificates (thereby exchanging claims against its own assets for claims against the CBCH), or it could raise new deposits from depositors who would be sold loan certificates. The bank would then use the proceeds to pay off the older depositors. In this way, the problem of bank-specific risk arising from informational asymmetries was resolved through a private system of coinsurance among banks.

Despite the efforts of the CBCHs to restrain member banks, they could not eliminate *all* moral hazard. Besides, there was the possibility of the CBCH itself being corrupted. Thus, there remained a role for monitoring by depositors. This, in turn, led to occasional runs on banks.

Reasons for Federal Deposit Insurance: Even though private arrangements can diminish the likelihood of bank runs, there are two reasons why they cannot eliminate them. First, even though a private arrangement like the CBCH provides depositors with some diversification, this diversification is limited by the size of the group of member banks. Size limitations may arise from transportation or information costs. Moreover, as the group grows larger, the cost to the CBCH of cheating by an individual bank diminishes, and the CBCH's incentive to monitor its members is weakened. This may be one reason why a large number of new clearinghouses sprang up within a 10-year period following the establishment of the New York CBCH in 1853, rather than a single "mega" clearinghouse emerging. A second weakness of

12. See Gorton and Mullineaux (1987).

private arrangements is that depositors can never be completely sure of the integrity of the arrangement. Thus, there was still some incentive for depositors to monitor the CBCH. In turn, this implies that panics could not be avoided.

The establishment of the Federal Reserve System in 1914 was partly in response to the inadequacy of private arrangements in performing key central bank functions. Nevertheless, the Fed could not prevent the banking panics of the Great Depression, and this eventually led to the establishment of federal deposit insurance. Two of the arguments for federal deposit insurance are discussed below.

(1) Money Supply: the Macroeconomic Argument: At a macroeconomic level, deposit insurance acts as a stabilizer by preventing reductions in the stock of money through bank failures.[13] Since commercial banks are the main providers of the nation's money stock, large-scale uninsured failures of commercial banks would reduce the national money supply. Deposit insurance helps to prevent this in two ways: (a) it replaces deposits that would otherwise be lost, and (b) it discourages banking panics by preserving public confidence.

The reason why deposit insurance has to be *federal* is the credibility of the federal government in its promise to meet all contractual payments. Because of its virtually unlimited authority to raise revenues through taxation, the federal government can meet payout commitments that may be far in excess of the deposit insurance fund. This taxation may be explicit (the government can simply raise taxes) or implicit (the government can print more money to repay depositors, thereby taxing by reducing the real value of each unit of money).

(2) Improving Consumer Welfare: the Microeconomic Argument: We have already noted the incentive of individual depositors to monitor the bank. This results in costly duplication of monitoring. In the numerical illustration of the previous section, the equilibrium involves all 30 vigilant depositors monitoring the bank even though monitoring by just one depositor would suffice. There are two ways in which federal deposit insurance helps to reduce overall monitoring costs. First, because a government agency (the federal insurer) is insuring deposits, the need of insured depositors to monitor is either eliminated (when deposit insurance is complete) or diminished (when deposit insurance is incomplete). Moreover, since the federal insurer must itself monitor banks, even uninsured depositors perceive a much smaller need to monitor. In other words, most of the monitoring burden is shifted from individual depositors to the federal insurer. This eliminates much of the duplicated monitoring encountered with uninsured deposits, *without* any residual monitoring incentives as with the private CBCH arrangement. Second, a federal deposit insurer can be expected to specialize in monitoring insured banks because it must deal with a large number of them. Thus, even apart from reducing duplication, there may be a direct reduction in monitoring costs. For example, in our numerical illustration, instead of monitoring costing 1 cent per audit, it might cost 3/4 cent per audit.

The overall effect of reduced monitoring costs will be to increase the *effective* interest rates on deposits,[14] but this benefit of deposit insurance may be offset by a host of implementation problems that we have yet to address.

13. See Scott and Mayer (1971).

14. To see this, imagine that in the previous numerical illustration, depositors can be assured that the bank will choose project B, and the total monitoring cost to ensure this choice is only 3/4 cent.

Banking Runs and Panics: Theories and the Empirical Evidence

Although the idea that deposit insurance can eliminate bank runs is an old one, research of the last decade has provided a clearer understanding of *why* bank runs and banking panics occur. In light of the recent S&L and banking turmoil, linked by many to federal deposit insurance, alternative arrangements deserve careful consideration. This subsection offers a perspective that should be useful in thinking about these issues.

When informational imperfections interfere with the functioning of a market, governmental intervention may be warranted. An example is Akerlof's lemons problem in the used car market (recall Chapter 1); "lemons laws" protect used car buyers in many states. Another example is the Federal Aviation Authority's regulation of airline safety and the Federal Drug Administration's regulation of the medicinal drug market. In these markets, it is very costly for consumers to let the market provide the necessary disciplining of providers. Similarly, if banking panics disrupt the productive sector of the economy, federal deposit insurance may be warranted if it is effective in reducing the likelihood of panics. The two main theories discussed below explain *how* deposit insurance can prevent runs and panics.

(a) The "Sunspots" Theory of Bank Runs: This theory maintains that bank runs are triggered by completely random events like "sunspots."[15] Suppose that we live in a two-period world with three points in time: $t = 0, 1, 2$. Individuals are risk averse. At $t = 0$, individuals have endowments of wealth that they wish to invest in projects. Each project requires a \$1 investment at $t = 1$, pays off \$R for sure at $t = 2$ if not liquidated earlier, and has positive NPV, that is, each offers a rate of return sufficiently higher than the riskless rate (which is zero) if continued until $t = 2$. Let $R > \$1$. However, if the project is liquidated prematurely at $t = 1$, then there is a loss of productive efficiency and the project pays off only \$1. At $t = 0$, individuals are unsure of their future preferences for the timing of their consumption. At $t = 1$, they receive a "preference shock" and learn whether they are about to die or will live another period. If they are about to die, they want to withdraw the money they have invested and consume it immediately at $t = 1$. If they learn that they will live, then they want to leave their money in the projects and consume \$R at $t = 2$. For the population as a whole, a (random) fraction, f, of individuals are "diers" at $t = 1$ and a fraction, $1 - f$, are "livers."[16]

What would happen without a bank? Well, if you discover at $t = 1$ that you are a dier, you will liquidate your investment and consume \$1. Call the first-period consumption C_1^D, that is, $C_1^D = 1$, and your second-period consumption, $C_2^D = 0$. If you discover that you are a "liver," then you will choose to consume nothing at $t = 1$ (that is, $C_1^L = 0$) and you will consume an amount $C_2^L = R$ at $t = 2$. Thus, the *nonbank outcome* is the pair $\{C_1^D = 1, C_2^D = 0\}$ or the pair $\{C_1^L = 0, C_2^L = R\}$, depending on the individual's type. Is this the best outcome

15. See Noyes (1909) and Gibbons (1968). Bryant (1980) and Diamond and Dybvig (1983) provide contemporary treatments. The discussion below is based on Diamond and Dybvig (1983).

16. The terms "diers" and "livers" are not meant to be taken literally, but merely represent those with preferences for immediate consumption (diers) and for deferred consumption (livers).

from the standpoint of an individual at $t = 0$? The answer is obviously no! Since you are a *risk-averse* individual, you would like some insurance at $t = 0$ against a random future shock to your own preference for consumption. This is where a bank can help.

The basic idea is as follows. To provide risk-averse individuals insurance against preference shocks, a bank can arise to promise those withdrawing at $t = 1$ a little more than \$1 and those withdrawing at $t = 2$ a little less than \$R, still ensuring that the promised payoff at $t = 2$ exceeds that at $t = 1$. Since $R > 1$, this is simply a temporal redistribution of the individual's wealth from a state of nature in which wealth is relatively high to one in which it is relatively low, that is, a classic insurance scheme. Compare this to a capital market that also redistributes temporally, but involves no insurance aspect. As long as the bank has a reasonably good idea of how many individuals will withdraw on average at $t = 1$ (this is similar to insurance companies estimating likely outcomes based on actuarial tables), it can structure the deposit contract in such a way that a known fraction of projects are liquidated at $t = 1$ to pay off the withdrawers. Note that more projects will need to be liquidated than there are withdrawers because each depositor is promised more than \$1 and the liquidation value of each project at $t = 1$ is \$1. Hence, those waiting until $t = 2$ will receive less than \$R. This is a nice arrangement because the $t = 2$ payoff exceeds the $t = 1$ payoff, so if a depositor can "afford" to wait until $t = 2$, he will. Thus, one possible outcome is that only the diers withdraw at $t = 1$ and all the livers wait until $t = 2$. All depositors are better off than they would be without a bank because they have received some insurance at $t = 0$ against unpredictable future changes in their preferences.

The fly in this ointment, however, is that the entire scheme rests delicately on the assumption that none of the livers withdraws at $t = 1$. But what if a liver believes that others like him might "panic" and withdraw at $t = 1$? If this belief is justified, it would be foolish for him to be the only patient depositor since the bank will have to liquidate *all* its projects at $t = 1$ and there will be nothing left to disburse at $t = 2$. So he will attempt to withdraw at $t = 1$ as well. In other words, the beliefs of the livers at $t = 1$ are crucial. If a representative liver believes others will withdraw at $t = 1$, he will too, and a panic run at $t = 1$ is a Nash equilibrium. On the other hand, if a representative liver believes others will wait until $t = 2$, he will too, and this is a Nash equilibrium as well. These beliefs are unrelated to the quality of the bank's assets.

How do you preclude the bad Nash equilibrium? One way is to provide deposit insurance. If the claims of all depositors are insured, then the livers know that they are guaranteed a payoff at $t = 2$ that is independent of the actions of other depositors. Hence, all livers will withdraw only at $t = 2$, and there will be no bank run. The example in the box below makes these ideas concrete.

Example 10.2 Suppose there are 100 risk-averse individuals, each with \$1 to invest in a project at $t = 0$. The project will yield \$1 if liquidated at $t = 1$ and \$2.25 if liquidated at $t = 2$. At $t = 0$, no individual knows what his "type" (denoting his consumption preference) will be at $t = 1$. If the individual turns out to be a "dier" (type D), then his utility function for consumption will be

$$U_D = \sqrt{C_1^D}.$$

If he turns out to be a "liver" (type L), then his utility function for consumption will be

$$U_L = 0.6\sqrt{C_1^L + C_2^L}.$$

These utility functions capture the idea that the dier benefits from consumption at $t = 1$ only, and the liver is indifferent between consuming at $t = 1$ or $t = 2$ (he gets equal utility from each) so that he will prefer the higher of the two consumptions. It is known at $t = 0$ that 40 percent of the individuals will end up being diers and 60 percent will be livers at $t = 1$. Compute the *ex ante* ($t = 0$) expected utility of each individual if (i) there is no bank and each individual invests in his own projects, and (ii) there is a bank that accepts a \$1 deposit from each individual and invests all the proceeds in 100 projects.

Solution We will solve this problem in six steps. First, we calculate each individual's expected utility absent banks. In this scheme, an individual receives \$1 if he consumes at $t = 1$ and \$2.25 if he consumes at $t = 2$. Second, we introduce a bank that is a mutual owned by the 100 depositors. It promises \$1.1 to each depositor withdrawing at $t = 1$ and \$2.1 each to those withdrawing at $t = 2$. Each depositor experiences a higher expected utility at $t = 0$ with this scheme than in the nonbank case. Third, we show that the intermediated outcome leads to a (good) Nash equilibrium in which all type-D depositors withdraw at $t = 1$ and all type-L depositors wait until $t = 2$. Fourth, we show that there is also a bad Nash equilibrium in which all depositors withdraw at $t = 1$. Fifth, we note that the bank run described in step 4 arises for no particular reason, but that it is possible whenever the existence of the bank makes depositors better off. Finally, in step 6 we show how deposit insurance can eliminate the Nash equilibrium.

Step 1 Consider first the nonintermediated situation. Let us assume, for simplicity, that the diers/livers fractions (0.4 and 0.6) can be viewed as subjective probability assessments of all individuals at $t = 0$. Then each individual believes that he faces a 0.4 chance of being of type-D at $t = 1$ and a 0.6 chance of being of type-L. In the nonintermediated case, $[C_1^D = 1, C_2^D = 0]$, and $[C_1^L = 0,\ C_2^L = R = \$2.25]$. Hence, each individual's expected utility will be

$$\begin{aligned} E(U) &= 0.4 \times \sqrt{1.0} + 0.6 \times 0.6 \times \sqrt{2.25} \\ &= 0.9400. \end{aligned}$$

Step 2 Now consider a bank, owned by its 100 depositors. It provides insurance against depositor preference shocks with a demand deposit offering $C_1^* > \$1$ and $C_2^* < \$R$ (where asterisks denote first- and second-period consumptions in the intermediated case), with the stipulation that C_1^* and C_2^* are mutually exclusive. For example, suppose the bank announces at $t = 0$ that $C_1^* = \$1.1$. Then, with 40 depositors withdrawing at $t = 1$, the bank will need to pay out \$44, and this requires premature liquidations of 44 projects. The remaining 56 projects will yield a total payoff of $56 \times \$2.25 = \$126$ at $t = 2$. The bank will be able to promise each of the 60 depositors withdrawing at $t = 2$ an amount $C_2^* = \$126/60 = \$2.1$. The expected utility of a depositor at $t = 0$ will be

(*Continued*)

$$\begin{aligned} E^*(U) &= 0.4 \times \sqrt{1.1} + 0.6 \times 0.6 \times \sqrt{2.1} \\ &= 0.9412. \end{aligned}$$

Hence, every individual is made better off by the bank that provides *consumption smoothing*.

Step 3 The step-2 outcome is a Nash equilibrium among depositors. Each type-D depositor's Nash equilibrium strategy is to withdraw at $t = 1$ since that gives him his highest utility (his utility from consumption at $t = 2$ is zero). If each type-L depositor *takes as given* the Nash equilibrium strategy of the *other* type-L depositors (to wait until $t = 2$ to withdraw), then no type-L depositor can do better by withdrawing at $t = 1$. This is because withdrawal at $t = 2$ gives a type-L a utility of

$$0.6 \times \sqrt{2.1} = 0.8695$$

whereas withdrawal at $t = 1$ gives a utility of

$$0.6 \times \sqrt{1.1} = 0.6293.$$

Thus, a Nash equilibrium is needed for all type-D depositors to withdraw at $t = 1$ and all type-L depositors to wait until $t = 2$.

Step 4 The "good" outcome is not the only Nash equilibrium, however. There is also a "bad" Nash equilibrium with a bank run. To see this, suppose that the representative type-L depositor believes that all the other type-L depositors will withdraw at $t = 1$ rather than $t = 2$.[1] What should you, as the "representative" type-L depositor, do?

Suppose you also decide to withdraw at $t = 1$. The bank will then observe that all 100 depositors wish to withdraw. All 100 projects will have to be liquidated to obtain \$100. According to the sequential service constraint, the bank will pay \$1.1 each to the first 90 depositors and the remaining \$1 to the 91st depositor; the last nine depositors receive nothing. If you wait until $t = 2$ to withdraw (when all the other depositors withdraw at $t = 1$), you get nothing. If you rush to the bank at $t = 1$, then assuming that your position in the queue is decided randomly (with equal probability of being at any position in the queue), you have a 0.9 probability of receiving \$1.1, a 0.01 probability of receiving \$1, and a 0.09 probability of receiving nothing. Clearly, your optimal strategy is to withdraw at $t = 1$ too. Thus, it is also a Nash equilibrium for *all* depositors to withdraw $t = 1$. This equilibrium is a *bank run*.

Step 5 Two points are noteworthy. First, the bank run in step 4 arises for no particular reason. We are not in a position to say which Nash equilibrium will arise. Hence, while we can say that a bank run is a possibility, we cannot say *why*. Second, a simple way for the bank to eliminate this type of run is to stipulate that withdrawers of demand deposits at $t = 1$ can receive only \$1. In this case, the bank does not need to liquidate more projects than there are withdrawers at $t = 1$, so that a depositor who waits until $t = 2$ will surely receive \$R, Thus, it is optimal for *every* type-L depositor to wait until $t = 2$, regardless of what the other type-L depositors do. But in this case the bank's demand deposit contract provides no risk sharing and the bank adds no value over the nonintermediated case. Hence, runs are a possibility whenever the bank adds value.[2]

Step 6 Deposit insurance can eliminate the bank run equilibrium *without* trivializing the bank. To see this, imagine that a governmental insurer were to guarantee that any individual withdrawing at $t = 1$ will receive \$1.1 and any individual withdrawing at $t = 2$ will receive \$2.1. Then, only the good Nash equilibrium survives.[3]

1. Do not ask why. This point is to see if this *can* be a Nash equilibrium. That is, conditional on such a belief about the behavior of others, does it pay for the representative type-L depositor to also behave like that?
2. You will note that the bank exists here for a different reason from that in Chapter 3.
3. Suspension of convertibility will work just as well. The bank could announce at $t = 0$ that only the *first* 40 withdrawers at $t = 1$ will be paid \$1.1 each. Remaining withdrawals can occur only at $t = 2$. This will do the trick, but only if the fraction of diers is known deterministically at $t = 0$. If this fraction is random, then the bank will not know *ex ante* when to suspend convertibility. In this case, deposit insurance is necessary to eliminate the bad Nash equilibrium without sacrificing the risk-sharing service banks.

The message of this theory is this: In the absence of deposit insurance, even a perfectly healthy bank faces the threat of a run, given the SSC associated with demand deposits. In other words, runs can result from shifts in the beliefs of individuals, unrelated to the "real" economy or the health of the banking system. Bank runs are simply random manifestations, a *force majeure* triggered even by "sunspots." In French, the term for a bank run is colloquially *sauve qui peut* (every man for himself).

Although some runs reflect sunspot phenomena, it is difficult to verify empirically what precisely triggered a run. Banking panics, on the other hand, have often been triggered by adverse information about banks. We now turn to an informational theory of bank runs.

(b) Adverse Information and Bank Runs: Suppose that we have three types of individuals.[17] We still have the diers (type-D individuals) who must consume at the end of the first period and represent a fraction, f, of all individuals. But among the livers, (type-L individuals), we now have a fraction who receives information about the terminal ($t = 2$) value of the bank's assets. In the previous theory, we assumed that this value, \$R, was nonrandom and known to everyone. Assume now that $\tilde{R}$ is a random variable with a commonly known expected value, R. Let $\tilde{R} = H > 0$ with probability p and $\tilde{R} = 0$ with probability $1 - p$. Thus, at $t = 0$, no individual knows either the $t = 2$ value of $\tilde{R}$ or what his type (D or L) will be at $t = 1$. However, at $t = 1$, each individual discovers whether he is a D or L, and some fraction, q, of the Ls also comes to know the value $\tilde{R}$ will take at $t = 2$. Nobody knows how many individuals of each type there are at $t = 1$ (that is, both the fraction f and the fraction q are random).

17. This discussion is based on Chari and Jagannathan (1988).

The choice problem faced by the Ds and the informed Ls at $t = 1$ is straightforward. All the Ds will line up to withdraw their deposits. If the informed Ls learn that $\tilde{R} = H$, then it is better for them to defer withdrawal until $t = 2$, thereby avoiding premature project liquidation. But if the informed Ls learn that $\tilde{R} = 0$, then it pays for them to withdraw whatever they can at $t = 1$.

Consider now the choice problem of the uninformed Ls. They can withdraw at $t = 1$ or wait until $t = 2$. Their decision will be based on their assessment of the $t = 2$ value of the bank's assets. Although they cannot directly observe this value, they can infer it by observing the length of the withdrawal queue at $t = 1$.[18] In drawing this inference, they realize that some people are in the withdrawal queue at $t = 1$ because they have discovered that they are Ds. But they do not know *how many* such individuals there are. This means that when they observe the length of the withdrawal queue at $t = 1$, they are unsure whether all are Ds or whether some are informed Ls.

It is true, however, that the longer the queue the more likely it is that it contains some informed Ls with adverse information about the bank. If the uninformed Ls knew for sure that the queue contained informed Ls, they would withdraw their money at $t = 1$, and if they knew for sure that it contained only Ds, they would defer withdrawal until $t = 2$. But when they cannot be sure, they use the queue length as a *noisy* signal of the information possessed by the informed Ls. Thus, they withdraw their deposits at $t = 1$ if the queue is sufficiently long, and they defer withdrawal until $t = 2$ if the queue is shorter.

Defining a bank run as a situation in which uninformed Ls withdraw at $t = 1$, we see that a bank run is more likely when some depositors receive adverse information about the bank. The reason is that as the informed Ls line up to withdraw their funds, they increase the queue length. This induces the uninformed Ls also to seek withdrawal of their deposits. Thus, a bank run results from depositors attempting to detect the bank's condition from the length of the withdrawal queue. However, since their learning is "noisy" (they occasionally confuse liquidity-motivated withdrawals with informed withdrawals), they make both type-I and type-II errors.[19] That is, they sometimes do not run the bank when they should (when the queue is relatively short but consists of informed Ls: a type-II error if the null hypothesis is that the bank is healthy); and they sometimes run the bank when they should not (when the queue length is relatively long but consists only of Ds: a type-I error). Because runs can sometimes occur when they should not, deposit insurance may improve welfare by eliminating the possibility that uninformed Ls will erroneously withdraw.

(c) The Empirical Evidence on Panics: Strictly speaking, neither of the two theories of bank runs discussed just above explains *panics*. According to the sunspots theory, a bank run is a completely random event, so there is no reason for a run to precipitate a panic, although a panic could come about by pure chance. According to the adverse information theory, a run is caused by information *specific* to a bank. Once again, there is no reason for a run to be contagious. These then are theories of bank runs and not banking panics.

The adverse information theory, however, can be adapted to provide an explanation for banking panics. Suppose there is information about some event that is relevant to the fortunes of all banks. That is, there is a systematic risk element that affects all banks.

18. This inference will usually be noisy. Formally, the inference may be made using Bayes' rule (see Chapter 1).

19. A type-I error in statistics is when the decision maker rejects the (null) hypothesis although it is true and a type-II error is when he accepts (or more appropriately, fails to reject) the null hypothesis although it is false.

Unlike the standard Capital Asset Pricing Model, however, assume that the systematic risk is *not* commonly known. Individuals may then attempt to infer something about systematic risk from their observations of presumably related events. For example, the failure of a large bank may cause depositors to believe that general economic conditions have deteriorated, and this may lead to a panic. The intuition is similar to that of the adverse information theory. According to that theory, depositors infer something about their bank from the behavior of fellow depositors. Here, depositors at one bank infer something about their bank from the behavior of depositors in *other* banks.

An example of an event that may reveal adverse systematic information is a recession, or a bank run during a recession. During the period from 1873 to 1914, every major business cycle downturn was accompanied by a banking panic.

Empirical evidence supports this version of the adverse information hypothesis. If banking panics are indeed systematic events, then there must be a change in the risk perceptions of individuals prior to a panic, and this, in turn, must cause a change in the deposit/currency ratio. That is, the perceived risk variable must achieve some critical value at the panic date. Also, the movements in the risk predictors and in perceived risk should occur at panic dates and not at other dates. If such movements occurred at other dates, then there should have been panics at those dates.

An empirical examination of panics in the pre-Federal Reserve era provides insight into the relationship between changes in risk perceptions and banking panics.[20] To serve as a proxy for perceived risk, empiricists use unanticipated changes in the liabilities of failed businesses.[21] This is reasonable since the fortunes of nonfinancial firms affect the fortunes of banks. As Table 10.2 shows, panic dates correspond to the timing of the largest values of the liabilities shocks. Panics also follow the business cycle peak by several months.

The study also indicates that the percentage change in the currency/deposit ratio is significantly correlated with the perceived risk measure. Thus, the data for the pre-Fed period support the notion of a threshold value of perceived risk that triggers panics. More recent research indicates that the banking panics during 1890–1909 were triggered by net movements of deposits away from the money-center banks and low levels of excess reserves. Changes in stock market values had little effect.[22]

The formation of the Federal Reserve System in 1914 and the initiation of deposit insurance in 1934 had a significant influence on the timing of panics. In the period from 1914 to 1933, we see from Table 10.3 that changes in the perceived risk measure were large enough in at least one instance (June 1920) to cause panics during the pre-Fed period, but resulted in no panics in the post-Fed period.

The introduction of deposit insurance again significantly changed depositor behavior. In the period from 1935 to 1972, until after deposit insurance was introduced, there were several instances of large failed business liabilities shocks, none of which resulted in panics. Thus, deposit insurance appears to have served its purpose.

Deposit Insurance Pricing and Moral Hazard

Until the 1980s, the pricing of federal deposit insurance was largely *risk insensitive*. That is, each bank was charged an insurance premium that depended only on its

20. The evidence reported here is from Gorton (1988).

21. In Gorton's (1988) empirical study, this variable was measured by the residuals ("error terms") from an estimated time-series model.

22. See McDill and Sheehan (2006).

TABLE 10.2 The Relationship Between the Timing of the Largest Unanticipated Changes in the Liabilities of Failed Businesses and the Timing of Banking Panics in the National Banking Era

NBER Chronology Peak-Trough (Business Cycle)	Timing of Largest Value of Unanticipated Changes in Liabilities of Failed Businesses	Panic Date
Oct. 1873–Mar. 1879	Dec. 1873	Dec. 1873
Mar. 1882–May 1885	June 1884	June 1884
Mar. 1887–Apr. 1888	Nov. 1887	No panic
July 1890–May 1891	Dec. 1890	Dec. 1890
Jan. 1893–June 1894	July 1893	July 1893
Dec. 1895–June 1897	Oct. 1896	Oct. 1896
June 1899–Dec. 1900		No panic
Sep. 1902–Aug. 1904		No panic
May 1907–June 1908	Feb. 1908	Dec. 1907
Jan. 1910–Jan. 1912	Mar. 1910	No panic
Jan. 1913–Dec. 1914	Mar. 1914	Sep. 1914

Source: Gorton, Gary, "Banking Panics and Business Cycles," *Oxford Economic Papers* 40, 1988, 751–781.

TABLE 10.3 The Relationship Between the Timing of the Largest Unanticipated Changes in the Liabilities of Failed Businesses and the Timing of Banking Panics in the Federal Reserve Era

Peak-Trough (Business Cycle)	Timing of Largest Value of Unanticipated Changes in Liabilities of Failed Businesses	Panic Date
Aug. 1918–Mar. 1919	Nov. 1918	No panic
Jan. 1920–July 1921	June 1920	No panic
May 1923–July 1924	Nov. 1923	No panic
Oct. 1926–Nov. 1927	Apr. 1927	No panic
Aug. 1929–Mar. 1933	Dec. 1929	Oct. 1930
		Mar. 1931
		Jan. 1933

"The change in perceived risk in June 1920 was large enough to have caused a panic in the pre-Fed Era."

Source: Gorton, Gary, "Banking Panics and Business Cycles," *Oxford Economic Papers* 40, 1988, 751–781.

volume of deposits, and not on its riskiness. Many have charged that this heightened incentives for insured depository institutions to take excessive levels of risk. Note that institutions like banks can increase risk in a variety of ways. However, for the purposes of this discussion, we will focus on the bank's incentive to invest in assets with high default risk. Although deposit insurance premiums are now risk sensitive, only a limited number of risk categories are used and at best, the premiums are only crudely related to risk for most banks. In this section, we will show how the imperfectly risk-sensitive structure of deposit insurance pricing also creates incentives for excessive risk-taking by banks.[23]

23. The discussion here is based on Merton (1977).

Deposit Insurance as an Option: Consider an insured bank (both principal and interest on deposits are insured) that has raised deposits requiring the bank to repay \$B at the end of the period. Let \$V be the total value of the bank's assets at the end of period. Now, if $V \geq B$, then the depositors receive \$B from the bank and the bank's shareholders receive \$(V − B). If $V < B$, then the bank fails. Its shareholders receive nothing, whereas the deposit insurer takes possession of the bank's assets and pays out \$B to the depositors. The net *loss* to the deposit insurer in this case is \$(B − V). Thus, the end-of-period payoffs to the different parties can be written as

Shareholders:	Max[0, V − B]
Depositors:	B
Deposit Insurer:	Min[0, V − B], which is either zero (when $V > B$) or negative (when $V < B$).

The effect of deposit insurance is to create an additional cash inflow to the firm of −Min[0, V − B] dollars. But −Min[0, V − B] can also be written as Max[0, B − V]. Hence, if G(T) is the value to the firm of the deposit insurance guarantee when the length of time remaining to maturity of the deposits is T, then on the date of maturity,

$$G(0) = \text{Max}[0, B - V]. \tag{10.1}$$

You will recall now from our discussions of options in Chapters 1 and 8 that the payoff structure in (10.1) is identical to that of a *put option* at expiration. To see this, imagine that V is the (random) value of the underlying security on which the option is written, and B is the exercise (or strike) price. Then, as the owner of the put, you will exercise your option to *sell* the security to the option writer at \$B if the value of the security, V, is less than B. In this case, your gain from exercising the option will be \$(B − V). On the other hand, if $B < V$, then you will let the option expire unexercised, and your gain will be zero.

The Cost of the Option: In other words, when the FDIC insures a bank's deposits, it is writing a put option in favor of the bank. The cost to the FDIC of providing this insurance is simply the value of the put option. We can calculate this value using the option pricing formula developed by Black and Scholes (1973):

$$G(T) = Be^{-rT}\Phi(x_2) - V\Phi(x_1) \tag{10.2}$$

where

$$x_1 \equiv \frac{\log(B/V) - \left[r + \frac{\sigma^2}{2}\right]T}{\sigma\sqrt{T}}$$

$$x_2 \equiv x_1 + \sigma\sqrt{T}.$$

Here r is the instantaneous risk-free interest rate, $\Phi(\bullet)$ is the standard normal cumulative distribution function, V is the current value of the bank's assets, and σ^2 is the variance rate per unit time of the logarithmic changes in the value of the assets. It is assumed that all the Black-Scholes assumptions are satisfied.

The Cost Per Dollar of Deposits: We can also compute the appropriate deposit insurance premium per dollar of deposits. If depositors are promised a repayment of \$B at a time, T, in the future, then the current value of these (riskless) deposits will be

$$D = Be^{-rT}. \tag{10.3}$$

Let $g = G(T)/D$ be the cost (to the FDIC) of the deposit insurance guarantee per dollar of insured deposits. Then, using (9.2) and (9.3) we can write

$$g(d,\tau) = \Phi(h_2) - \frac{1}{d}\Phi(h_1) \tag{10.4}$$

$$\text{where } h_1 \equiv \frac{\left[\log - \frac{\tau}{2}\right]}{\sqrt{\tau}} \tag{10.5}$$

$$h_2 \equiv h_1 + \sqrt{\tau}. \tag{10.6}$$

Here $d \equiv D/V$ is the current deposit-to-asset value ratio for the bank, and $\tau \equiv \sigma^2 T$ is the variance of the logarithmic change in the value of the assets during the term of the deposits.

Properties of a Risk-Sensitive Deposit Insurance Pricing Scheme: A few points are worth noting. First, an increase in the deposit-to-asset value ratio causes an increase in the cost per dollar of deposit insurance to the FDIC, that is,

$$\partial g/\partial d = \Phi(h_1)/d^2 > 0.$$

Similarly, as τ increases, so does the cost of deposit insurance, that is,

$$\partial g/\partial \tau = \Phi'(h_1)/2d\ \sqrt{T} > 0.$$

Here the prime denotes a derivative; hence, $\Phi'(h_1)$ is the standard normal density function at h_1. This is a well-known property of options; their value increases with the volatility of the underlying security. Hence, the FDIC should charge a higher deposit insurance premium for banks with lower capital-to-total-assets ratios and higher volatility in the value of total assets. Alternatively, in a regime in which the FDIC charges each bank a fixed premium per dollar of insured deposits, rather than g (which is a function of d and τ), banks with higher capital ratios and lower asset risks *subsidize* those with lower capital ratios and higher asset risks, assuming that the FDIC breaks even on average.

The Option Feature and Moral Hazard: These observations also highlight the moral hazard inherent in deposit insurance. Since g is the *value* to the bank of deposit insurance per dollar of insured deposits, a bank can increase this value by reducing its capital and increasing its asset volatility. To the extent that the premium charged is insensitive to these initiatives of the bank, a *shareholder-wealth-maximizing* bank has an incentive to increase financial leverage and asset volatility. *Figure 10.1* illustrates this incentive graphically.

In Figure 10.1, the curve AB is the total expected return on the bank's assets, net of bankruptcy costs.[24] This curve peaks at σ^*. The expected return to depositors, as represented by the straight line CD, remains constant because we assume that deposits are completely insured. The total expected return to depositors and the FDIC is equal to the deposit yield plus the deposit insurance premium minus the expected bankruptcy costs. This total *expected* return, represented by the curve BF, is constant for $\sigma < \sigma_f$ (some threshold variability) because the probability of bankruptcy is zero in this range. Then, as the probability of bankruptcy rises, the total expected return to the FDIC and the depositors declines. Since the depositors are completely protected and the deposit insurance premium is insensitive to σ, it is the expected return to the FDIC that falls very rapidly as σ arises. Consequently, even though the total expected return on the bank's assets is falling as σ increases beyond σ^*, the expected return to the bank's shareholders is increasing in this range. In fact, the shareholders' expected return peaks at $\sigma_m > \sigma^*$.

The optimal level of risk (as represented by σ) depends on the decision maker's objective. If the objective is to minimize the liability of the deposit insurer, then the optimal risk choice is $\sigma = \sigma_f$. If the objective is to maximize the bank's total expected return on assets, then the optimal risk choice is $\sigma = \sigma^*$. But if decisions are made to maximize the wealth of shareholders, the optimal risk choice is $\sigma = \sigma_m$. Thus, if the "socially desired" risk choice is σ^*, the bank will take more risk than the social optimum by choosing σ_m. This is the moral hazard of deposit insurance.

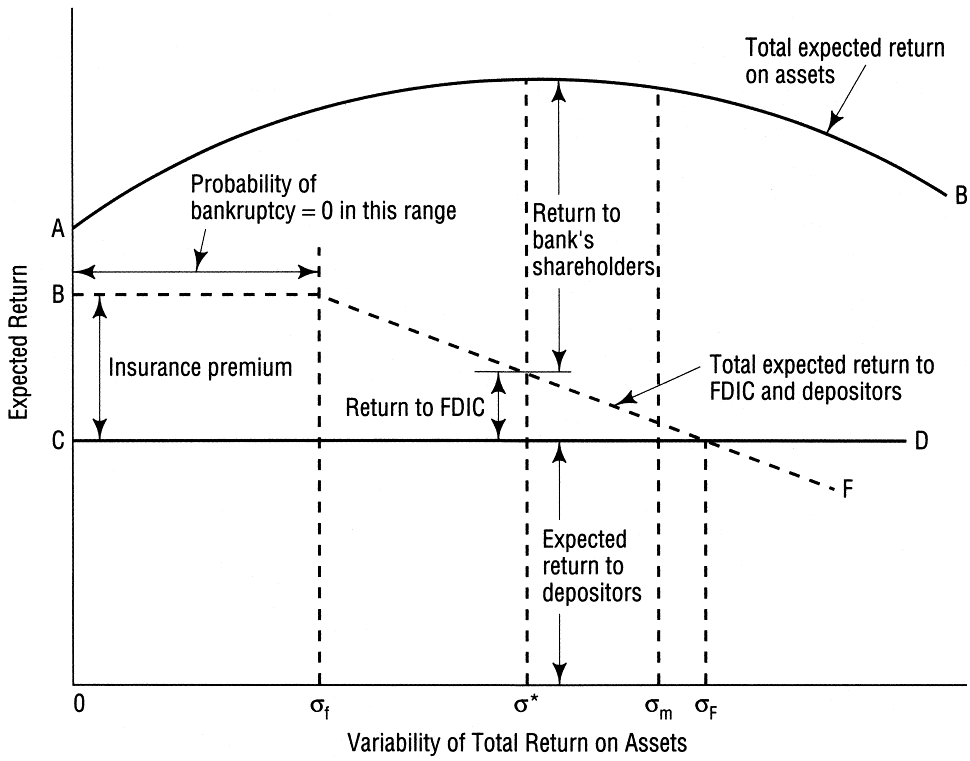


FIGURE 10.1 The Relationship Between Expected Return and Risk

24. Figure 10.1 is based on Keeton (1984), who provides a similar figure (see p. 32 in that paper).

Why the Concern With Moral Hazard in Banking?: You will recall from Chapters 5 and 6 that a similar moral hazard exists for nonfinancial firms that borrow from banks. However, with nonfinancial firms, the costs of this moral hazard are borne *ex post* by private lenders, who pass along these costs *ex ante* (through the pricing mechanism) to the borrower. Thus, the moral hazard gets priced among the contracting parties in equilibrium. In the case of banks and other federally insured depository institutions, however, these costs are borne *ex post* by the FDIC, and hence, eventually by the taxpayers. Of course, if the FDIC breaks even in aggregate, then these costs are passed along *ex ante* to the banking industry as a whole, and there is simply a redistribution of wealth across banks. That is, less risky banks end up subsidizing their riskier counterparts, with no direct wealth consequences for the taxpayers.

This analysis, as well as our earlier discussion of the similarity between a deposit insurance guarantee and a put option, indicates a role for safety regulation in banking. Given deposit insurance, banks have a propensity to lower capital and increase risk. Capital requirements and asset portfolio restrictions seek to address these distorted incentives arising from deposit insurance. However, the implementation of these regulatory devices has not always been effective.

In the box below, we provide an illustration of the effect of moral hazard in the context of the put option pricing formula.

Example 10.3 Consider a bank with federally insured deposits maturing in one year. Imagine that the bank's asset value changes monthly and you have been provided the following data on asset values for the past seven months (you may assume that the probability distribution of asset value changes remains stationary through time).

Month	Bank Asset Value (in millions of dollars)
1	100
2	101
3	99
4	102
5	100
6	98
7	97.605074

Suppose the bank's current deposit-to-asset value ratio is 0.95. Compute the value to the bank of the deposit insurance guarantee per dollar of insured deposits. Also compute the value of this guarantee for a higher deposit-to-total-asset-value ratio (of your choice), holding fixed the variance of asset value changes, and the value of this guarantee for a higher variance, holding fixed the deposit-to-total-asset-value ratio.

Solution We solve this problem in three steps. First, we will compute τ, the variance of bank asset values. Second, we compute h_1 and h_2 using the value of τ obtained in the previous step. Finally, in step 3 we calculate the cost of deposit insurance per dollar of insured deposits.

Step 1 To compute τ, we define V_t as the asset value in month t and V_{t-1} as the asset value in month $t - 1$. Thus, when we write the asset value in month 2, for example, we will write V_2, and when we write the ratio V_t/V_{t-1} in month 2, we will write V_2/V_1. We can construct the following table.

TABLE 10.4 Calculation of Asset Value Variance

A Month	B Asset Value V_t	C V_t/V_{t-1}	D $\log(V_t/V_{t-1})$	E D-Sample Mean	F $(E)^2$
1	100	—	—	—	—
2	101	1.01	0.00995	0.013988	0.0001957
3	99	0.9802	−0.02000	−0.015962	0.0002548
4	102	1.0303	0.02985	0.033888	0.0011484
5	100	0.9804	−0.01979	−0.015752	0.0002481
6	98	0.9800	−0.0202	−0.016162	0.0002612
7	97.605074	0.9959701	−0.004038	0	0

In this table we compute the "sample mean" by adding up the entries in column D and dividing by 6 to obtain −0.004038. Column E is then obtained by subtracting the sample mean from each entry in column D. Column F is merely each entry in column E squared. Now,

$$\sigma^2 = \frac{\text{sum of all entries in column F}}{5}$$
$$= \frac{0.0021082}{5} = 0.0004216.$$

Note that we divide by 5 because we lose one degree of freedom in computing the variance. Now, $\tau = \sigma^2 T = 0.0004216 \times 12 = 0.005$ approximately. Note that $T = 12$ since the deposit maturity is 1 year and asset values change monthly.

Step 2 Next, we compute h_1 using (10.5) as

$$h_1 = \frac{\log(0.95) - (0.005/2)}{\sqrt{0.005}}$$
$$= -0.76076$$

and h_2 using (10.6) is

$$h_2 = -0.76076 + \sqrt{0.005} = -0.69005.$$

Step 3 Using (10.4), we can now compute g as

$$g = \Phi(-0.69005) - \frac{1}{0.95}\Phi(-0.76076) \cong 0.0099.$$

Thus, the value to the bank of having the deposit insurance guarantee is roughly 99 cents per \$100 of insured deposits. This is much higher than the current premium of approximately 25 cents per \$100 of insured deposits. In Table 10.5, we present calculations for a variety of deposit-to-asset value ratios and values of τ. Note that if we increase d to 1 and hold τ fixed at 0.005, the value of g rises to \$2.82 per \$100 of insured deposits. This illustrates the bank's incentive for leverage emanating from deposit insurance. Similarly, if we hold d fixed at 0.95 and increase τ to 0.006, the

(*Continued*)

value of g rises to $1.209 per $100 of insured deposits. This illustrates the bank's incentive to take on more risky assets.

TABLE 10.5 Cost of Deposit Insurance per Dollar of Insured Deposits

Cost of Deposit of Insurance (g)	Deposit-to-Asset Value Ratio	Variance (τ)
0.00055	0.85	0.00600
0.00040	0.85	0.00550
0.00028	0.85	0.00500
0.00018	0.85	0.00450
0.00011	0.85	0.00400
0.00326	0.90	0.00600
0.00274	0.90	0.00550
0.00223	0.90	0.00500
0.00176	0.90	0.00450
0.00132	0.90	0.00400
0.00093	0.90	0.00350
0.00060	0.90	0.00300
0.00015	0.90	0.00200
0.01209	0.95	0.00600
0.01102	0.95	0.00550
0.00992	0.95	0.00500
0.00880	0.95	0.00450
0.00765	0.95	0.00400
0.00647	0.95	0.00350
0.00528	0.95	0.00300
0.00287	0.95	0.00200
0.00172	0.95	0.00150
0.00072	0.95	0.00100
0.00033	0.95	0.00075
0.03089	1.00	0.00600
0.02958	1.00	0.00550
0.02820	1.00	0.00500
0.02676	1.00	0.00450
0.02523	1.00	0.00400
0.02360	1.00	0.00350
0.02185	1.00	0.00300
0.01784	1.00	0.00200
0.01545	1.00	0.00150
0.01262	1.00	0.00100
0.01093	1.00	0.00075
0.00892	1.00	0.00050
0.00631	1.00	0.00025
0.00564	1.00	0.00020
0.00489	1.00	0.00015
0.00399	1.00	0.00010
0.00282	1.00	0.00005
0.00126	1.00	0.00001

Source: Merton, Robert C., "The Cost of Deposit Insurance and Loan Guarantees," *Journal of Banking and Finance* 1, June 1977, 10.

The option pricing approach indicates factors that must be considered in setting the deposit insurance premium. The premium per dollar of insured deposits must be sensitive to the volatility of the bank's assets and to its deposit-to-total-asset ratio. If not, the bank will have an incentive to reduce its capital and increase its asset risk in the interests of its shareholders. The option pricing approach is not meant to be taken literally as a precise way to set the deposit insurance premium, since many of the standard Black-Scholes assumptions are not satisfied.[25] For example, the asset values of banks often exhibit jumps rather than following a continuous path through time as assumed by Black-Scholes. In any case, the numerical values in Table 10.5 suggest the magnitude of the gains to banks from exploiting risk-insensitive deposit insurance pricing.

Empirical Evidence on Moral Hazard

Apart from the anecdotal evidence on moral hazard, there is now substantial scientific evidence to support the theories we have reviewed. Federal deposit insurance has been in existence for banks since 1934, but the more visible problems were encountered only during 1970–90. This suggests that there must have been *countervailing forces* in the past that diminished the risk-taking propensity created by deposit insurance. The empirical evidence we discuss here sheds some light on these forces.

(a) Some Evidence on the Effect of Federal Deposit Insurance on Risk-Taking: The Case of Credit Unions: Federal deposit insurance was extended to credit unions in 1971 when the National Credit Union Administration (NCUA) was formed, and the coverage limits are currently the same as for banks and thrifts.

Credit unions: (i) make loans to their own members, (ii) make loans to other credit unions, and (iii) engage in loan participations with other credit unions. A credit union's asset portfolio consists primarily of: (i) secured loans for the purchase of consumer durables, and (ii) investments in low-risk assets such as government bonds, loans to other credit unions, and deposits with commercial banks.

A credit union can increase its risk by decreasing its capital cushion and by increasing the fraction of its total assets invested in high-risk assets. An empirical examination provided support for this hypothesis.[26] *Figure 10.2* is a graph illustrating the behavior of capital ratios (defined as capital divided by total assets) for federal credit unions over the 1949–1992 period. In 1970, just prior to the adoption of federal deposit insurance, the capital ratio was about the same as in 1949. There was a slight decline during the transition period from the preinsurance regime to the insurance regime. The sharpest decline occurred during the insurance period. This is consistent with the prediction of a moral hazard associated with deposit insurance.

(b) The Relationship Between Market Power in Banking and Moral Hazard: As mentioned earlier, a major puzzle is why the deposit insurance system in the United States worked so well for so many years despite the risk-taking incentives provided by federal deposit insurance and why problems surfaced only recently. One explanation is that a bank's risk-taking propensity depends on the value of its charter. The higher the charter value—the capitalized value of its future cash flows—the weaker is the

25. In addition, it may be difficult to ensure that the deposit insurer measures bank risk without error. See Flannery (1991) for a discussion of the implications.

26. See Clair (1984).

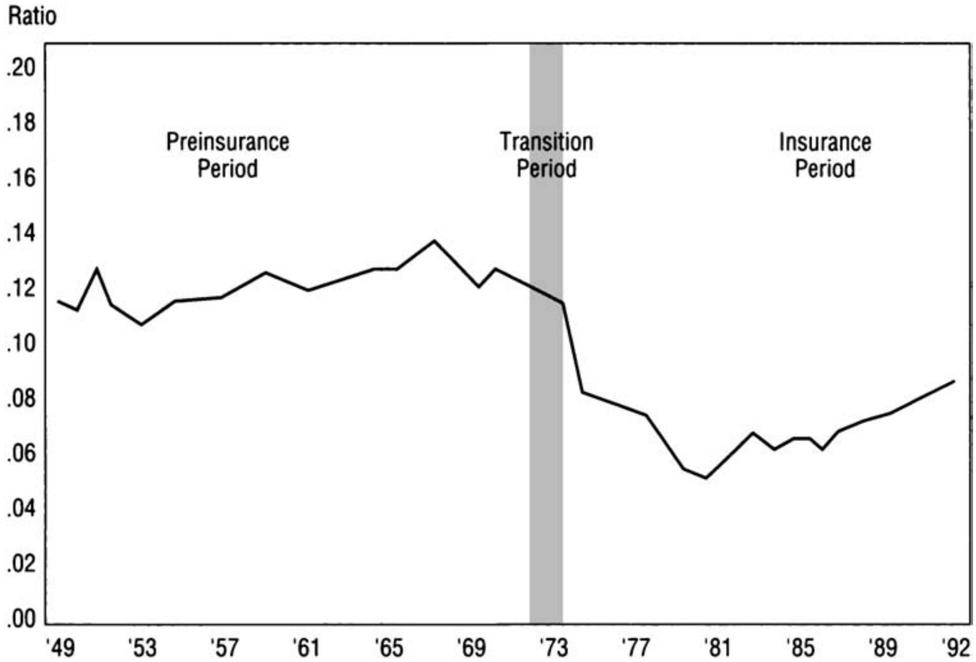


FIGURE 10.2 Capital Ratio for Federal Credit Unions
Source: National Credit Union Administration.

bank's incentive to take risk. This is because higher risk implies a higher likelihood of insolvency, in which case the insurer takes possession of the bank, and the charter is lost. Thus, the higher the value of this charter, the greater is the bankruptcy cost for the bank. In the past, various anticompetitive restrictions gave banks market power that enhanced the value of charters. The loss to the bank from losing its charter in the event of bankruptcy provided a counterbalance to the incentive for excessive risk-taking due to fixed-rate deposit insurance.[27] The deregulation that took place in the 1980s increased banking competition but lowered the value of bank charters. Greater risk-taking was predictable.

Evidence supports this theoretical prediction. *Figure 10.3* is a graph of the time series behavior of the average capital/total assets ratio of the 25 largest bank holding companies in the United States from 1952 to 1986. The decline in this ratio is significant.

A *direct* test of the relationship between risk-taking and charter value would need to have some measure of the capitalized value of future rents, or market power. One such measure is "Tobin's q," which is approximated as the ratio of the market value of assets (market value of common equity plus the book value of liabilities) to the book value of assets. The higher the q ratio, the larger is the charter value, relative to the book value of its assets. Since bank risk-taking is also not directly observable, a proxy is needed. A reasonable proxy is the interest cost on large uninsured CDs. The

27. For the theory, see Chan, Greenbaum, and Thakor (1992). The empirical evidence discussed below is from Keeley (1990).

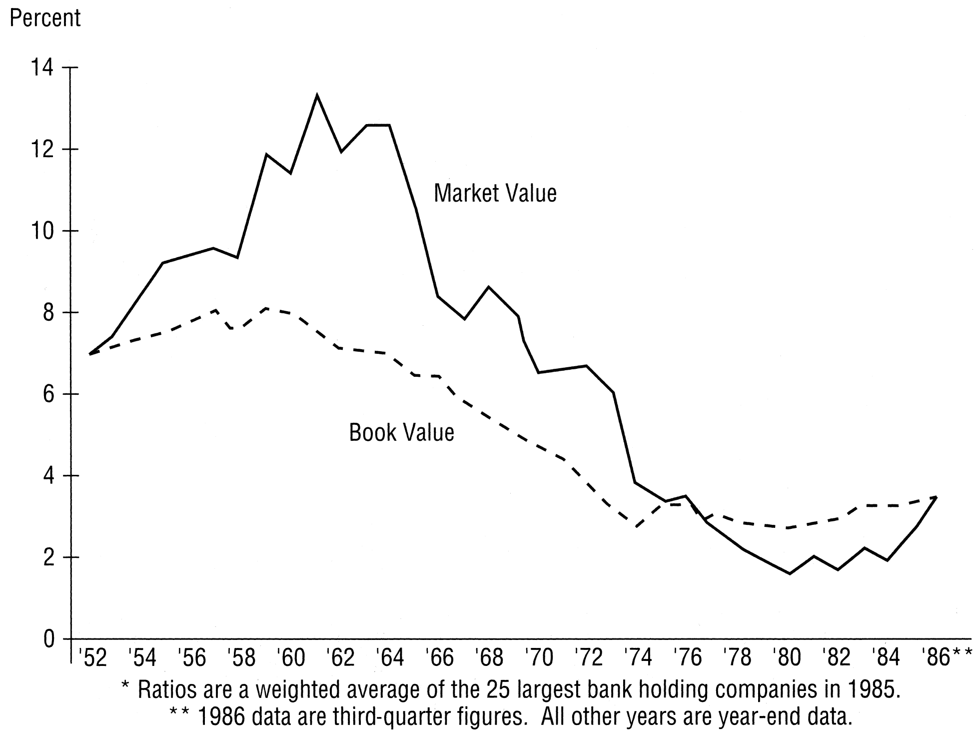


FIGURE 10.3 Capital-to-Asset Ratios, Market and Book Values
Source: Keeley, Michael C., "Deposit Insurance, Risk, and Market Power in Banking," *American Economic Review* 80, December 1990, 1183–1200.

holders of such CDs should be sensitive to the bank's riskiness and demand higher interest rates from riskier banks. The evidence is quite compelling. Each 1 percent increase in the q ratio results in a 16 to 18 basis point reduction in the average CD cost. Moreover, this relationship is statistically significant. Thus, bank risk-taking appears to have increased substantially in the 1980s owing to deregulation that diminished bank charter values.

To provide a comparison with more recent data, we have provided in *Figure 10.4* the capital ratios in book and market value terms for the 25 largest bank holding companies from 1959–2005. The effect of the capital regulation that began with the Basel I Accord is evident, as capital ratios exhibit an upward drift beginning in the late 1980s.

We also show the capital ratios during 1992–2005 in *Figure 10.5*. This figure shows that capital ratios have remained relatively flat during this time in book value terms, and above the minimum requirement of 4 percent for Tier-1 capital. The large increase in market-value-based capital ratios during 1994–2000 is probably a reflection of the high levels of the overall stock market during that time.

In *Figure 10.6,* we show the number of bank failures during 1992–2004. As is evident, the number of failures decreased dramatically during the early 1990s and has stayed relatively low as banks have operated with healthy capital ratios.

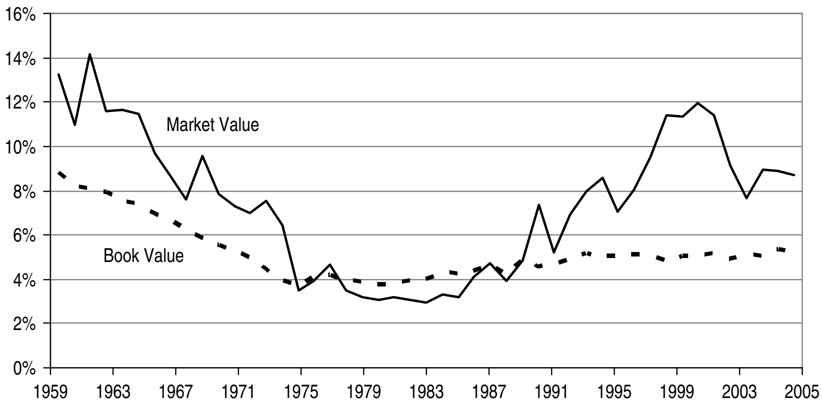


FIGURE 10.4 Capitalization of the 25 Largest Bank Holding Companies During 1959–2005
This Figure shows the weighted average capitalization of the 25 largest bank holding companies from 1959–2005. Capital ratios are expressed in book values (shareholder equity/total assets) and market values (market value of equity/market value of assets).
Source: Compustat and own calculations.

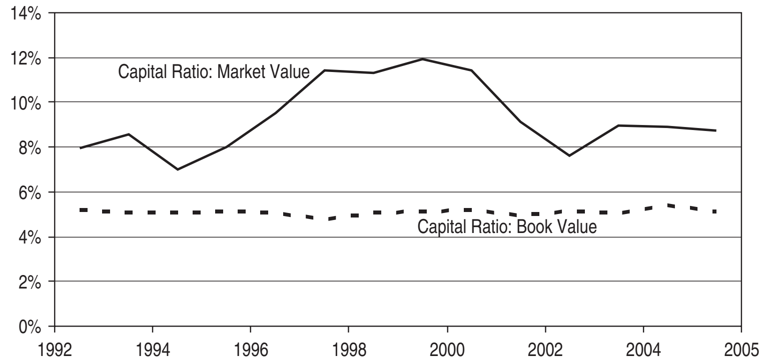


FIGURE 10.5 Capitalization of the 25 Largest Bank Holding Companies During 1992–2005
Source: Compustat and own calculations.

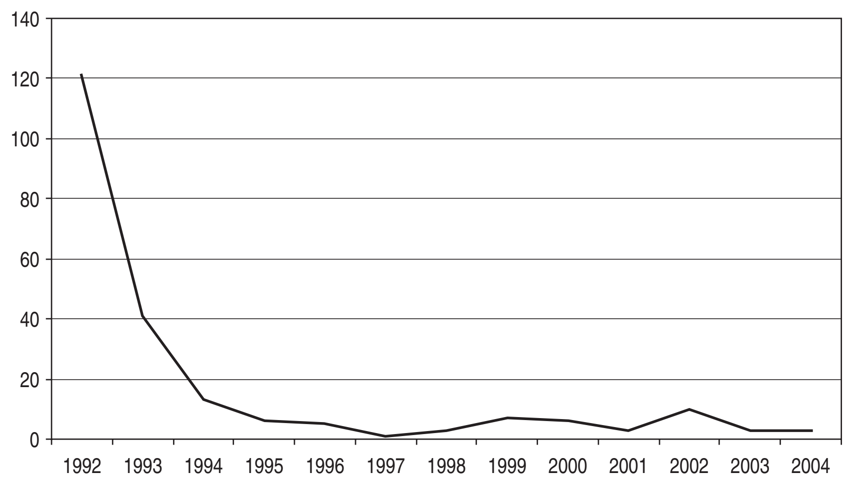


FIGURE 10.6 Bank Failures 1992–2004
Source: FDIC Historical Statistics on Banking: Closings and Assistance Transactions.

The Great Deposit Insurance Debacle

General Background

We have now reviewed both the theory and some empirical evidence about the effects of deposit insurance on depository institutions' risk-taking behavior. In trying to understand the great deposit insurance debacle of the 1980s, it is important to remember that until the mid-1970s deposit insurance worked remarkably well. But, two developments undermined federal deposit insurance. One is the lowering of bank charter values, which increased managers' incentives to take more asset risk and to also engage in fraud. The other is the decline in regulatory vigilance over the same period; this simply exacerbated the moral hazard problem of federally insured depository institutions.

The waste that resulted from the collapse of the thrift industry and the many banking failures in the 1980s can be classified into three categories: excessive risk-taking, excessive consumption of perquisites by top executives, and outright fraud. Moreover, these diversions/destructions of wealth were possible due to three factors working in concert: deposit insurance with risk-insensitive pricing, low charter values due to deregulation, and lax monitoring by regulators. This laxity in monitoring, caused by a lowered commitment of resources to supervision, was also compounded by cozy relationships between some regulators and the institutions they were supposed to be watching over. In *Figure 10.7,* we have provided a simple schematic to summarize these effects.

It is not as if S&L and bank managers woke up one morning in the 1980s and decided to change the way they made decisions in order to "rip off" the taxpayers.

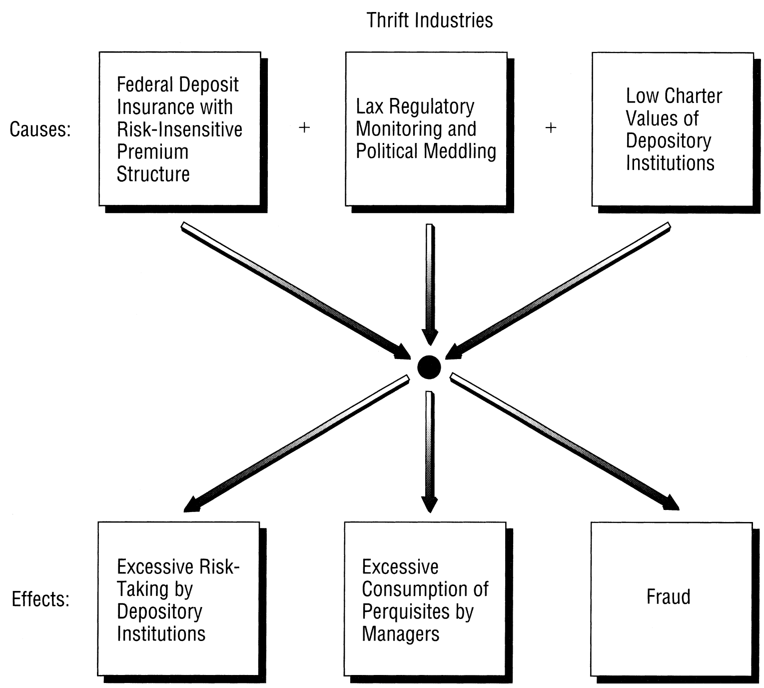


FIGURE 10.7 Schematic of Effects Responsible for Problems in Banking and Thrift Industries

The point is that their *incentives* were altered. Their *decision rule* was still the same, but the altered incentives changed their behavior. The reasons for the deposit insurance crisis can therefore be traced not just to the managers of depository institutions, but also to the politicians and regulators who pursued myopic and hasty policies. In what follows, we briefly discuss the causes and effects depicted in *Figure 10.7*.

Regulatory and Political Culpability

For some years, S&L regulators tried to ignore problems in the thrift industry. Hoping that problems would improve, regulators permitted insolvent institutions to continue to operate. Our analysis of credit risk in Chapters 5 and 6 highlighted the important incentive effects of capital on a borrower's risk-taking propensity. The same is true of depository institutions; their propensity to take risk is greater when capital is lower. When capital is negative, excessive risk-taking is easy to predict. Indeed, the Federal Home Loan Bank Board (FHLBB) was quite aware that the thrift industry was in deep trouble in 1981, but chose not to close all the insolvent institutions.[28]

This inaction was part of a broader regulatory malaise. The main findings of a 2-month *USA Today*-Gannett News Service investigation are listed below.[29]

- Some regulators had close ties to the industry.[30]
- In some cases, *regulators* suggested that S&Ls try to grow rapidly and to invest in risky ventures as a way of quickly boosting profits.
- Regulatory agencies lacked the powers and the human resources to monitor rapidly growing S&Ls.
- Congress repeatedly refused requests to add S&L examiners, and told the FHLBB to go easy on problem S&Ls.[31]

Excessive Risk-Taking

As previously discussed, the three prominent ways to detect excessive risk-taking involve examining capital-to-total-asset ratios, interest rates on large (uninsured) CDs, and the assets in the institution's portfolio. The last two are briefly discussed below.

28. A 1990 issue of *The American Banker* quoted Mr. Richard T. Pratt, then chairman of the FHLBB, "Had we liquidated the S&L industry in 1981, it would have cost $178 billion–$380 billion in today's dollars. It would have been the most foolish public policy that could have possibly been undertaken."

29. See *USA TODAY*, February 14, 1989.

30. Mr. Tom Huston, former Iowa state banking superintendent, claims that regulators traveled too much at industry expense. He said, "They were so loved and so well-treated ... that no wonder they couldn't make a rational decision."

31. Mr. Edwin Gray, FHLBB chairman from May 1983 to June 1987, blames Congress and the Reagan administration for failing to give regulators more power, and he blames the powerful S&L lobby for influencing them. In *USA TODAY* (2/14/1989), Mr. Gray was quoted as saying the following: "We were asking Congress and the Reagan administration for help and getting nothing. We had a rag-tag bank of 700 examiners, who were expected to monitor $1 trillion in assets and 3,300 S&Ls. Sometimes our examiners were hired away by the S&Ls they were examining." It turns out that entry-level examiners were paid $14,000 per year during this time, and the turnover rate was 25 percent.

Higher Interest Rates on CDs: Riskier institutions must pay higher interest rates on large CDs, or conversely, those institutions that offer to pay higher interest rates on their deposits anticipate investing in high-risk, high-yield assets to cover their deposit funding costs. Since risk-taking incentives are the strongest in insolvent and nearly insolvent institutions, one would expect such institutions to be paying the highest rates. This is precisely what happened in the southwestern United States, where the S&L industry was devastated. Higher interest rates offered by insolvent institutions led to a self-fulfilling prophecy. When a depository institution has low net worth, it is expected to invest in riskier assets, so that depositors demand relatively high interest rates. These high interest rates, in turn, increase the attractiveness of high-yield, high-risk assets to the institution, thus completing the cycle. Moreover, in order to compete with insolvent institutions, solvent institutions may be compelled to offer higher interest rates on their deposits, leading to stronger incentives to invest in riskier assets.

Investments in High-Risk Assets: There is ample evidence of excessively risky investments by S&Ls. These investments included loans to developers to build ski resorts, speculative positions in government securities, junk bond portfolios, and so on.

Excessive Consumption of Perquisites by Managers

Although it is empirically difficult to determine whether a given level of perquisites consumption by a manager is "appropriate," some of the examples are striking and suggestive of abuse. These include institutional purchases of planes to transport top managers from their places of residence to their offices, payments for escort services, offices lined with expensive antiques and paintings, and gold-painted toilets. Many of the institutions where such apparent abuse occurred were investigated by the FSLIC.

Fraud

Estimates of *direct* losses to the government due to fraud by S&L managers range from $8 billion to $15 billion, and fraud is suspected in 80 percent of failed S&Ls. Parties, mansions, airplanes, women, Rolls-Royces, and Cayman Island bank accounts are some of the perks that S&L executives showered upon themselves as they looted federally insured deposits.

The S&L crooks also caused failures of S&Ls run by honest managers, by selling them stakes in their bad loans. For example, the now-insolvent First Federal Savings and Loan of Malvern, Arkansas, bought an interest in a doomed $44 million loan to a high-rise condo in Honolulu, which subsequently defaulted.

Many of the fraud cases are very complex. Shady S&Ls and equally shady borrowers combined dozens of loans, companies, and properties into convoluted deals to cover personal use of S&L deposits. Some S&Ls made borrowers pay big one-time fees—4 percent to 10 percent of the loan—in order to obtain loans. The S&Ls would report these fees as income, which boosted profits. Many loans were never repaid, leaving the property in the S&L's hands. An S&L executive might get a kickback for participating in the scheme. In Texas, this strategy was described as: "Heads, I win. Tails, FSLIC loses."

Following the S&L debacle, the government has filed approximately 100,000 civil suits against S&L executives, directors, owners, borrowers, and others believed responsible for contributing to the insolvency of S&Ls. The success of these prosecution efforts, and of attempts to recover some of the losses due to fraud, negligence, and simple mismanagement, remains uncertain.

To summarize, the greatest banking debacle since the Great Depression was *not* just an "unfortunate break" or an outcome of exogenous changes in the banking environment. Increasing competition increased interest-rate volatility and deregulation reduced the profitability of depository institutions, substantially diminishing charter values. Models of bank behavior predict increased risk-taking by federally insured institutions in such a setting, suggesting a need for improved regulatory monitoring. Unfortunately, safety was sacrificed at the same time that the industry was deregulated, as resources devoted to regulatory supervision were decreased. Regulatory ineptness and political meddling compounded the effects of poorly-thought-out initiatives.[32]

Banking Fragility, Deposit Insurance and Developments Since the Great Deposit Insurance Debacle

We have seen in this chapter that deposit insurance induces moral hazard and invites banks to engage in reckless risk-taking. That is, there is an inherent paradox in the use of deposit insurance as a way to diminish the likelihood of bank runs and banking fragility. The safer banks feel due to deposit insurance, the greater is their risk-taking propensity! It is for this reason that it may be socially efficient to impose a limit on the level of deposit insurance, thereby leaving room for market discipline, which then opens up the possibility of bank runs and banking fragility. In other words, there may be an "optimal" amount of banking fragility that strikes the right balance between the market discipline associated with the possibility of bank runs to temper banks' risk-taking incentives and the need to ensure that the likelihood of runs is not so high as to make banking excessively fragile.[33]

One could argue that one way to cope with deposit insurance–related moral hazard is to use capital requirements as an instrument to reduce banks' proclivity to take excessive risk. Regulatory reforms associated with the Basel I and Basel II capital accords and FDICIA of 1991 (see the next two chapters) lend strong support to the hypothesis that sufficiently high capital requirements can be effective in controlling risk.[34] The incidence of bank failures during the 1990s and 2000–2005 has been remarkably low, and the FDIC has been building up its reserves.

The real question, however, is whether we need deposit insurance in the first place, for a lot of the regulatory apparatus we observe would be unnecessary were it

32. We recommend reading Adams (1990), Mayer (1990), and White (1991) for accounts of the many factors that contributed to the implosion of the thrift industry in the United States.

33. This implication can be drawn from Calomiris and Kahn (1991). It is a point that has been made by Diamond and Rajan (2001). For other analyses of banking fragility, see Allen and Gale (2001).

34. Coval and Thakor (2005) explain that a certain level of bank capital may even serve the purpose of ensuring the viability of a bank that seeks to serve as a "bridge" between borrowers and savers with divergent beliefs.

not for deposit insurance.[35] But would we not have excessive bank runs without deposit insurance? This is actually an open question. Mutual funds have no deposit insurance and we have not observed any runs. At the end of the day, a fundamentally sound banking system, backed up by a credible lender of last resort, may not be as fragile without deposit insurance today as it may have been in the past. If this is true, the entire system of deposit insurance and regulations may have to be reconsidered.

Conclusion

We have devoted this chapter to an extensive discussion of the deposit contract, liability management, and deposit insurance. The nature of the deposit insurance contract is such that it leaves the bank vulnerable to runs, and the banking system vulnerable to panics. It appears that deposit insurance served its purpose of minimizing bank runs and panics. Indeed, for almost 50 years since the inception of federal deposit insurance in 1933, failure rates in the banking and thrift industries have been abnormally low compared to other industries. Moreover, this stable environment meant that liability management was not a pressing issue for banks.

But all that changed in the 1970s and 1980s. As interest-rate volatility increased and interest-rate restrictions were relaxed and then eliminated, liability management became a significant concern for banks. Moreover, a combination of deregulation, heightened volatility in market prices, lax regulatory monitoring, political interference, and corrupt executives in federally insured institutions significantly undermined the safety of the industry, and imposed monstrous losses on the deposit insurance funds. It is somewhat ironic that these events were quite predictable, in light of what was known *prior* to these events. Regulatory reforms that followed have helped to significantly improve banking stability.

Review Questions

1. What are the main economic features of the demand deposit contract and how do these features discipline management when deposits are uninsured?
2. What measures were used to cope with bank runs and panics prior to federal deposit insurance? Why were these not entirely satisfactory?
3. What is a bank run and how can you explain a run on *economic* grounds?
4. How does deposit insurance prevent runs and panics?
5. Explain the similarity between deposit insurance and a common stock put option and how this leads to moral hazard.
6. Why did deposit insurance work so well in the United States until 1980 despite the obvious moral hazard, and why did it fail after that?
7. Discuss the roles of bank managers, accountants, regulators, and politicians in the "great banking/S&L debacle."
8. What is liability management and what are its main objectives?
9. What is the agency problem between the shareholders and managers of a bank in liability management?

35. See Miller (1995) for a forceful argument in favor of dismantling federal deposit insurance.

10. Is moral hazard unethical, illegal, or neither? Can you outline a conceptual framework for defining unethical behavior by a depository institution?
11. What aspects of S&L/bank behavior would you consider unethical? For example, were junk bond investments unethical? Be sure to take an *ex ante* perspective and not use "20-20 hindsight."
12. Why do you think unethical behavior became so rampant in the last decade and not prior to that? Did people change? Did morals decline in general? Did the environment change? Can you relate unethical behavior during this period to similar behavior during other periods in history?
13. Consider a bank that receives a \$1 deposit from each of 200 different depositors at $t = 0$. It invests \$25 of shareholders' equity in the bank and lends \$200, keeping \$25 as cash reserves. Out of the 200 depositors, there are 75 depositors (called type-D_1 depositors) who are capable of monitoring the bank's management; the remaining depositors (called type-D_2 depositors) have kept their money in the bank simply for transactions and safekeeping. The cost of monitoring the bank for an individual type-D_1 depositor is \$0.03 per period.

    The bank has two mutually exclusive investment opportunities. Project (or loan) A pays \$300 with probability 0.6 and zero with probability 0.4 at $t = 1$. Project B pays \$250 with probability 0.8 and \$220 with probability 0.2 at $t = 1$. If the bank chooses one of these two projects, the probability that the bank will actually end up with that project is 0.7. With probability 0.3, the bank will have inadvertently chosen the other project. Thus, we assume that the bank may make errors in project choice. By monitoring the bank, a type-D_1 depositor can discover the bank's true project choice at some point in time intermediate between $t = 0$ and $t = 1$, say at $t = 1/2$. These depositors can, if they desire, force liquidation of the bank by withdrawing their deposits at $t = 1/2$. Note that the bank's loans/projects mature at $t = 1$. If they are liquidated at $t = 1/2$, they are worth only \$70 to the bank. Under the terms of the deposit contract, the bank promises to pay 15 percent interest (conditional on the bank having the financial capacity to do so) if deposit withdrawal occurs at $t = 1$, and no interest if withdrawal occurs before that. Thus, a depositor is entitled to \$1.15 if she withdraws at $t = 1$, and \$1 if she withdraws at $t = 1/2$. The risk-free discount rate is zero and all agents are risk neutral.

    All the type-D_2 depositors plan to withdraw at $t = 1$, but each is subject to a random liquidity-motivated desire to withdraw at $t = 1/2$. To simplify, we will assume that even though no one knows in advance which (type-D_2) depositors will wish to withdraw at $t = 1/2$, the fraction of those who will wish to withdraw is known to be 25/125. That is, 25 type-D_2 depositors will wish to withdraw at $t = 1/2$. Assume that the bank's managers make decisions in the best interests of their shareholders. Compute the equilibrium strategies for the bank and its depositors.

References

Adams, James Ring, *The Big Fix: Inside the S&L Scandal*, John Wiley & Sons, Inc., New York, 1990.

Allen, Franklin, and Douglas Gale, *Comparing Financial Systems*, Cambridge, MA: MIT Press, 2001.

Association of Reserve City Bankers, *Capital Issues in Banking*, Chicago, December 1988.

Black, Fisher, and Myron Scholes, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy* 81, 1973, 637–659.

Bryant, John, "A Model of Reserves, Bank Runs, and Deposit Insurance," *Journal of Banking and Finance* 4, 1980, 335–344.

Calomiris, Charles W., and Charles M. Kahn, "The Role of Demandable Debt in Structuring Optimal Banking Arrangements," *American Economic Review* 81–3, June 1991, 497–513.

Calomiris, Charles W., Charles M. Kahn, and Stefan Krasa, "Optimal Contingent Bank Liquidation Under Moral Hazard," unpublished manuscript, April 1991.

Chan, Yuk-Shee, Stuart I. Greenbaum, and Anjan V. Thakor, "Is Fairly Priced Deposit Insurance Possible?" *Journal of Finance* 47–1, March 1992, 227–246.

Chari, Varadarajan V., and Ravi Jagannathan, "Banking Panics, Information, and Rational Expectations Equilibrium," *Journal of Finance* 43, July 1988, 749–761.

Clair, Robert T., "Deposit Insurance, Moral Hazard, and Credit Unions," *Economic Review*, Federal Reserve Bank of Dallas, July 1984, 1–8.

Coval, Josh, and Anjan V. Thakor, "Financial Intermediation as a Beliefs-Bridge Between Optimists and Pessimists," *Journal of Financial Economics* 75–3, March 2005, 535–570.

Diamond, Douglas, and Raghuram Rajan, "Liquidity Risk, Liquidity Creation and Financial Fragility: A Theory of Banking," *Journal of Political Economy*, April 2001.

Diamond, Douglas W., and Philip Dybvig, "Bank Runs, Deposit Insurance, and Liquidity," *Journal of Political Economy* 91, June 1983, 401–419.

Flannery, Mark, "Debt Maturity and the Deadweight Cost of Leverage: Optimally Financing Banking Firms," working paper, University of Florida, 1992.

——, "Pricing Deposit Insurance When the Insurer Measures Bank Risk with Error," *Journal of Banking and Finance* 15, 1991, 975–998.

Gibbons, J.S., *The Banks of New York, Their Dealers, the Clearing House, and the Panic of 1857*, Greenwood Press (reprint of 1859 original), New York, 1968.

Gorton, Gary, "Banking Panics and Business Cycles," *Oxford Economic Papers* 40, 1988, 751–781.

Gorton, Gary, and Donald J. Mullineaux, "The Joint Production of Confidence: Endogenous Regulation and Nineteenth Century Commercial-Bank Clearinghouses," *Journal of Money, Credit and Banking* 19–4, November 1987, 457–468.

Hutchison, David E., and George G. Pennacchi, "A Framework for Estimating the Value and Interest Rate Risk of Retail Bank Deposits," working paper 92–30, Federal Reserve Bank of Chicago, December 1992.

Jacklin, Charles, "Banks and Risk-Sharing: Risk-Sharing: Instabilities and Coordination," in S. Bhattacharya and G.M. Constantinides (eds.), *Financial Markets and Incomplete Information*, Rowman and Littlefield, Totowa, N.J., 1989.

——, "Demand Deposits, Trading Restrictions, and Risk Sharing," in E.C. Prescott and N. Wallace (eds.), *Contractual Arrangements for Intertemporal Trade*, University of Minnesota Press, 1987.

Jacklin, Charles, and Sudipto Bhattacharya, "Distinguishing Panics and Information-Based Bank Runs: Welfare and Policy Implications," *Journal of Political Economy 96*, June 1988, 568–592.

Kahane, Yehuda, "Capital Adequacy and the Regulation of Financial Intermediaries," *Journal of Banking and Finance* 1, June 1977, 207–218.

Keeley, Michael C., "Deposit Insurance, Risk, and Market Power in Banking," *American Economic Review* 80, December 1990, 1183–1200.

Keeton, William R., "Deposit Insurance and the Regulation of Deposit Rates," *Economic Review*, Federal Reserve Bank of Kansas City, April 1984, 28–46.

Mayer, Martin, *The Greatest-Ever Bank Robbery*, Charles Scribner's Sons, New York, 1990.

McDill, Kathleen M., and Kevin P. Sheehan, "Sources of Historical Banking Panics: A Markov Switching Approach," FDIC Working Paper 2006–01, November 2006.

Merton, Robert C., "An Analytic Derivation of the Cost of Deposit Insurance Loan Guarantees," *Journal of Banking and Finance* 1, June 1977, 3–11.

——, "On the Application of the Continuous-Time Theory of Finance to Financial Intermediation and Insurance," *The Geneva Papers on Risk and Insurance* 14, July 1989, 225–262.

——, "Operation and Regulation in Financial Intermediation: A Functional Perspective," in P. Englund (ed.), *Operation and Regulation of Financial Markets*, Economic Council, Stockholm, 1993, 17–68.

Miller, Merton, "Do the M&M Propositions Apply to Banks?" *Journal of Banking and Finance* 19, 1995, 483–489.

Noyes, A.D., *Forty Years of American Finance*, G.P. Putnam's and Sons, New York, 1909.

Scott, Kenneth E., and Thomas Mayer, "Risk and Regulation in Banking: Some Proposals for Federal Deposit Insurance Reform," *Stanford Law Review* 23, May 1971, 857–902.

White, Lawrence, *The S&L Debacle*, Oxford University Press, New York, 1991.

PART ♦ VI

Bank Regulation

This page intentionally left blank

CHAPTER ♦ 11

Objectives of Bank Regulation

"There have been three great inventions since the beginning of time: fire, the wheel, and central banking."

Will Rogers

Glossary of Terms

BHC: Bank Holding Company.

GAAP: Generally Accepted Accounting Principles.

SEC: Securities and Exchange Commission, a regulatory body for capital markets in the United States.

NOW: Negotiated Orders of Withdrawal, a deposit account.

OECD: Organization for Economic Cooperation and Development.

Federal Funds Market: A market in which banks borrow and lend their cash-asset reserves for short durations.

OCC: Office of the Comptroller of the Currency, an agency within the U.S. Treasury Department that charters and regulates national banks.

FDIC: Federal Deposit Insurance Corporation.

OTS: Office of Thrift Supervision, a U.S. government agency within the U.S. Treasury Department with responsibility for regulating federally insured savings and loan associations.

Basis Point: One hundredth of a percent.

LLR: Lender of Last Resort.

LBO: Leveraged Buyout, the highly leveraged purchase of a firm.

LDC: Loans to Developing Countries.

FNMA: Federal National Mortgage Association.

GNMA: Government National Mortgage Association.

FHLMC: Federal Home Loan Mortgage Corporation.

Introduction

In the previous chapter, we discussed one aspect of bank regulation, namely deposit insurance. In this chapter, we discuss bank regulation in a more general context, with attention to the objectives and incentive effects of regulation.

Public regulation of banking has a long and checkered history, with roots extending back to sovereigns who reserved to themselves the rights of coinage. By impressing their imprimatur on a flat piece of metal, a coin would be struck that would trade at a premium to the metal's intrinsic value. The premium derived from the monetary services provided by the coin, which in turn was enabled by the coin's managed scarcity and the authenticity of the coin's ingredients signaled by the imprimatur. The premium or monopoly profit earned from the coin, called seigniorage, was appropriated by the sovereign, owner of the imprimatur. Seigniorage was one of the more efficient modes of taxation.

Like coins, bank deposits provide monetary services, and they too are artificially scarce, even if their ingredients are altogether ephemeral. They also generate seigniorage, and this makes banks an obvious target of government regulation. But public regulation has encouraged circumventing adaptations on the part of the banks, which in turn has led to ever more intrusive forms of regulation. This dynamic, sometimes referred to as the regulatory dialectic, has led to more encompassing regulation.

With the growth in bank regulation, numerous rationales for regulation have been advanced including:

- fostering competition,
- protecting institutional safety and soundness,
- consumer protection,
- credit allocation, and
- monetary control.

The manifestations of government regulation include disclosure requirements, antidiscrimination restrictions, community reinvestment standards, cash-asset reserve requirements, minimum capital requirements, branching and bank holding company restrictions, asset proscriptions, loans-to-one-borrower limitations, and deposit interest-rate ceilings, among others. Figure 11.1 summarizes the major objectives of regulation.

In the next section we explain the most basic reason for bank regulation, that arising from the governmental safety net. This is followed by a description of the agencies responsible for bank regulation. Governmental regulation of bank market structure and competition are examined next. Subsequent sections examine regulations

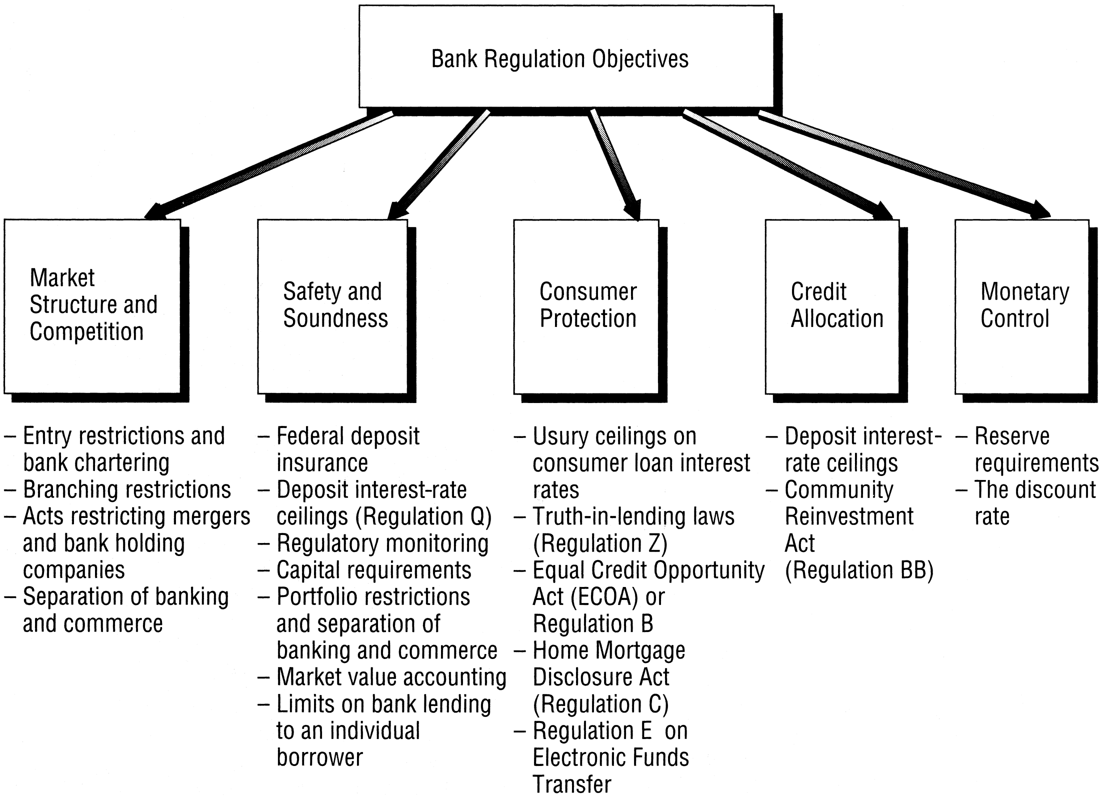


FIGURE 11.1 Objectives of Regulation

directed at safety and soundness, consumer protection, credit allocation, and monetary control. The history of regulation and regulatory reform are covered in Chapter 12.

The Essence of Bank Regulation

The provision of a governmental "safety net" for banks in the form of a lender-of-last-resort (LLR) facility, deposit insurance, or other guarantees creates the possibility of exploitation of the government by the banks.[1] This moral hazard necessitates a response if the safety net is to remain viable, and the response is normally some form of public regulation.

The Primitive Banker Without a Governmental Safety Net

Recall the goldsmith of Chapter 3 who evolved into a fractional reserve banker by printing warehouse receipts in excess of his or her gold holdings. The earnings on loans in that setting were seigniorage, a monopoly profit on the production of money

1. See, for example, Buser, Chen, Kane (1981) and von Thadden (2004). The literature dealing with the economics of bank regulation is reviewed by Bhattacharya, Boot and Thakor (1997).

in the form of receipts.[2] Now consider this same monopolist goldsmith confronted by a nascent central bank. Suppose the goldsmith's stock of gold has been purchased by the central bank in exchange for noninterest-bearing demand deposits at the central bank, and the central bank has in turn sold the gold in order to purchase interest-bearing government securities.[3] Balance sheets might appear as follows:

Central Bank

Government Securities 150	Deposits (Goldsmith's) 150

Goldsmith

Deposits at Central Bank 150	Receipts 1,000
Loans 850	

First, notice that the 150 in deposits at the central bank are the voluntary cash asset reserves of the goldsmith. These are optimal in light of the goldsmith's decision to circulate receipts of 850 in excess of liquid asset holdings. Second, notice that supporting the 1,000 in goldsmith receipt liabilities is 1,000 in collateral, 850 in the form of earning assets, and 150 in deposits at the central bank.

Thus, the seigniorage on the receipts money is shared between the goldsmith in the form of lower earnings and the central bank in the form of government securities earnings. The sharing is determined, to a first approximation, by the goldsmith's decision as to how many excess receipts to circulate by making loans. This decision presumably reflects the goldsmith's own risk-taking preferences. Notably, the central bank does *not* dictate a minimum fraction of deposits to be held at the central bank to back the goldsmith's receipt liabilities. Hence, there is no legal cash-asset reserve requirement.

The Governmental Safety Net and Moral Hazard

Now, in recognition of the goldsmith's vulnerability to bank runs (see Chapter 10), let us have the central bank institute an LLR facility. For concreteness, assume that the central bank stands ready to lend without limit to the goldsmith against performing (but illiquid) loan collateral at an interest rate of say 1 or 2 percent above the risk-free rate. Let us further assume that the volume of goldsmith receipts remains fixed. We can think of the volume of receipts as the monetary policy indicator, and the central bank can do *open market operations* (purchases and sales of government securities) to offset the bank's possible inclination to expand receipts as a result of the introduction of the LLR facility.[4]

2. Others could enter the business, of course, and drive down the profit, but even the first entrant could expand until the fear of withdrawals called a halt. In any case, owing to consideration of withdrawal risk, proliferation of receipts can be expected to terminate before the pressure on spreads has eliminated all seigniorage.

3. This would require that the goldsmith's customers accept deposits at other institutions or currency in lieu of the gold originally deposited. But this is not quite as bizarre as it might at first appear. U.S. bank deposits were (not easily) redeemable in gold until 1972 when President Nixon severed the last official link between gold and dollars. Few today seem to be aware or concerned. When President Reagan was elected in 1980, he appointed a commission to study the possibility of a return to the gold standard, but the idea soon died.

4. The assumption of a fixed volume of receipts is for simplicity only and does not compromise the basic argument.

Now the question is: How does the introduction of the LLR facility affect the bank's choice of reserve ratio, or to put it differently, how does the introduction of the LLR affect the bank's decision as to how many loans to make? Clearly, since the LLR represents an additional source of liquidity for the goldsmith, the need to hold nonearning central bank deposits is reduced.

The goldsmith/bank will therefore increase its lending, which will temporarily increase the volume of receipts outstanding. The central bank will feel compelled to sell government securities in order to restore the amount of receipts to 1,000.[5]

If we look at the new balance sheets, after the introduction of the LLR, we might find the following:

Central Bank

Government Securities	Deposits (Goldsmith's)
100	100

Goldsmith

Deposits at Central Bank	Receipts
100	1,000
Loans	
900	

Thus, the introduction of the LLR facility resulted in 50 of earning assets being shifted from the central bank to the goldsmith. This redistribution of earning assets is tantamount to a transfer of seigniorage from the government to the privately owned fractional-reserve goldsmith/bank, and is symptomatic of the moral hazard inherent in the LLR facility. The central bank provides the privately owned banks with a new layer of protection, the LLR facility, and because of its more secure position the goldsmith/bank sheds some of its own protection (cash assets) in order to expand earnings. But the expanded earnings of the goldsmith come, at least partly, at the expense of the central bank. From the central bank's viewpoint, this is clearly an unintended and exploitative side-effect of the LLR facility. And if carried far enough, all of the goldsmith's withdrawal risk will be transferred to the central bank. A private-sector risk of banking will have been nationalized, and all seigniorage will be transferred to the goldsmiths/banks.

Regulatory Response to Moral Hazard

This moral hazard threatens the viability of the LLR, and it therefore evokes an adaptive response by the central bank in the form of restrictions on bank behavior, such as legal cash-asset reserve requirements. These, together with sanctions for their violation, have clear analogs in private contracting. A fire insurance policy typically reduces the vigilance of the insured and thereby shifts additional risk to the insurer. The insurer reacts by requiring that the insured maintain minimum safety standards, and violations void the insurance coverage. In the case of deposit-taking banks, a large part of public regulation can be explained as protective responses to the moral hazards arising from safety-net provisions provided by the government.

5. The central bank's sale of its U.S. government securities extinguishes its deposit liabilities and thereby reduces the goldsmith's liquid assets. This prompts the goldsmith to reduce loans until the original 1,000 in receipts is re-established.

The most important among these are the LLR facility, deposit insurance, protection of the payments system, and the too-big-to-fail policy.[6] All of these create moral hazards that shift costs and risks from the private banks to the public (central bank) and, therefore, elicit restrictions on bank behavior designed to limit such exploitation.

The key of this theory of bank regulation is that the potential for deregulation is bound up with the span of the safety net. Deregulation, beyond the elimination of redundancies, requires *pari passu* shrinkage of the safety net that prompted the regulation. This nexus is inescapable, and deregulation rhetoric that ignores the trade-off is just that.

The Agencies of Bank Regulation

Paralleling the fragmented structure of the financial services industry is a similarly fragmented collection of public regulatory agencies. Each major fragment of the industry has its own dedicated regulatory agency, often duplicated at state and federal levels. For example, commercial banks, thrifts, and credit unions can be chartered (licensed) at either the state or the federal level. Therefore, each state will have a governmental agency charged with licensing, examining, supervising, and regulating thrifts, credit unions, and commercial banks. Likewise, the federal government licenses, regulates, supervises, and insures each deposit-taker. Life and casualty insurance companies are regulated principally at the state level, but recent failures, with effects that spilled across state borders, have evoked calls for more coordination of insurance regulation at the federal level. Even in the securities business, where the Securities and Exchange Commission dominates, corporations must seek approval of the states in which they are incorporated when issuing equity or debt securities.[7]

In this section, we discuss the agencies that regulate commercial banks. The complexity and fragmentation should be clear, even though we address only a smallish slice of the financial services industry.

Dual banking in the United States means that a bank can be licensed by either the federal government (the OCC), or by the state in which it is domiciled.[8] Even state-chartered commercial banks are likely to be regulated by at least two federal bank regulatory agencies since they are required to satisfy the Federal Reserve's cash-asset reserve requirements, and they are almost universally insured by the Federal Deposit Insurance Corporation. National banks are subject to regulation by three or more federal agencies. A distinction is typically drawn between regulation, the setting of rules, and supervision, which is monitoring compliance. The latter subsumes examinations and related activities. All of the federal bank regulatory agencies—the OCC, the Federal Reserve, and the FDIC, as well as the state banking agencies—have both regulatory and supervisory responsibilities.[9]

6. A concept closely related to "too-big-to-fail" is "too-many-to-fail." When too many banks are likely to fail, the regulator is inclined to bail out a lot of them. See Acharya and Yorulmazer (2007).

7. Futures have their own dedicated federal regulatory agency, the Commodities Futures Trading Commission (CFTC).

8. Federally chartered banks are required to have "national" in their names, and state-chartered banks are prohibited from including "national" or N.A. (national association) in their names. Likewise, federally chartered thrifts are required to have "federal" in their names whereas their state-chartered counter parts are prohibited from using "federal" in their name.

9. For some regulations, however, a regulator may not be a supervisor. For example, the Federal Reserve sets truth-in-savings regulations for all deposit-taking institutions, but supervises only banks.

Office of the Comptroller of the Currency (OCC)

Lodged within the Treasury Department, the OCC is the oldest existing federal bank regulatory agency. It was created pursuant to the National Bank Act of 1864 for the purpose of chartering and regulating "national banks." At the time, virtually all banks in the United States were state chartered and state regulated. Indeed, the federal government had been out of the business of regulating banks since the early 1830s when President Andrew Jackson stifled the rechartering of the Second Bank of the United States.[10]

Recall that 1864 was the time of the Civil War, and a national banking system was seen as an opportunity to provide a marker for Union bonds that were being sold in record amounts to finance the hostilities. So the National Bank Act imposed a 5 percent tax on the liabilities of state-chartered banks and commenced the chartering of national banks, which were required to hold U.S. government securities to satisfy reserve requirements. Deposits were far less important than bank notes at the time both as a means of payment and as a financing instrument of banks. The OCC's primary responsibility was the chartering and supervision of national banks.[11] National banks are a minority among commercial banks, but they tend to be the larger banks.

The Federal System

The Federal Reserve System was established in 1913 following a searching investigation prompted by a particularly disruptive financial panic in 1907. The principal purpose of the Federal Reserve was to provide LLR services to a banking system vulnerable to liquidity crises.[12] As LLR, the Federal Reserve stands ready to lend for short periods to liquidity-strapped banks against eligible collateral. So from the very outset, the Federal Reserve used cash-asset reserve requirements to deter banks from substituting the liquidity of the LLR facility for previously held cash assets.

All national banks were required to be members of the Federal Reserve System, whereas membership was voluntary for state-chartered banks. Only in 1980 were all insured commercial banks compelled to meet the cash-asset reserve requirements of the Federal Reserve.

The Federal Reserve has been a remarkably successful regulatory agency in at least two senses. It has managed to remain relatively free of scandal and it has, not coincidentally, enormously expanded the scope of its regulatory turf. Perhaps most important in the latter regard were the Bank Holding Company Act of 1956 and the Douglas Amendments thereto of 1970. These laws gave the Federal Reserve regulatory control over all bank holding companies. Practically every important bank in the United States is owned by a bank holding company, and the Federal Reserve has immense discretionary power over virtually every initiative taken by banks via their holding companies.[13] These typically include acquisitions and mergers both within

10. For further reading on the early history of U.S. banking, see Lash (1987).

11. The name of the agency traces to its intervention in bank note issuance.

12. The Federal Reserve was also charged with providing a flexible currency, and this led to the Federal Reserve note in use.

13. The regulatory discretion derives from the vagueness of the legislative mandate. For example, in evaluating nonbank acquisitions by bank holding companies, the Federal Reserve is instructed to judge whether the contemplated acquisition is so closely related to banking as to be a "proper incident thereto." Some decry such vagueness as the hallmark of bad legislation. Others laud the ambiguity as a possible benefit in mitigating moral hazard problems. For more on this issue, see Boot and Greenbaum (1992) and Boot, Greenbaum, and Thakor (1993).

and outside of banking. Explicit permission is required from the Federal Reserve for each and every subsidiary formed or purchased by a bank holding company, and the criteria for approval leave broad scope for discretion by the Federal Reserve. No wonder the chairman of the Federal Reserve is sometimes described as the second most powerful person in the United States.

In addition to regulating bank holding company activities and determining cash-asset reserve requirements, the "discount rate" (LLR), ceilings on deposit interest rates (this power has been substantially restricted since 1986), and capital requirements, the Federal Reserve also has vast powers in the realm of consumer protection. These powers were legislated by the Community Reinvestment Act of 1977, the Home Mortgage Disclosure Act of 1975, the Truth in Savings Act of 1991, as well as a raft of other laws. This legislation gave the Federal Reserve the power to combat discrimination in lending and employment, failure to disclose information relevant to banking and other transactions, and the failure of banks to provide credit in their local communities. The authority for ensuring compliance with these regulations is shared by the Federal Reserve and other regulators.

The Federal Reserve's vast regulatory discretion together with its power to influence capital markets places it first among the many public regulators of financial institutions.[14]

The Federal Deposit Insurance Corporation (FDIC)

The youngest of the three federal banking agencies, the FDIC, was established in the depths of the Great Depression in 1933. As with its predecessors, there was profound political ambivalence about government deposit insurance. Despite a collapsing banking industry and a similarly compelling need to restore public trust, President Roosevelt spoke out strongly against federal deposit insurance for the very reasons we now use to explain recent S&L and bank losses. The patrician Roosevelt was a half-century ahead of his time. Nevertheless, he ultimately signed the Banking Act of 1933 that provided federal insurance of bank deposits for the first time in U.S. history, and with it came the FDIC.[15] The original plan called for protections of the first $2,500 in each bank account, but over the years coverage has expanded 40-fold.[16]

In addition to the explicit insurance, the FDIC has often remunerated depositors with balances in excess of the stated limit so that very few "uninsured" depositors since 1933 have lost money owing to bank or S&L failures. This second layer of implicit insurance coverage is provided at the discretion of the FDIC and is typically rationalized in one of two ways. It is either less costly to compensate all rather than some, or imposing losses on some is too destabilizing to the financial system and the failed bank. This is of value since liquidation adds costs to compensate uninsured depositors and preserve the bank as an ongoing entity, possibly as a part of another bank.

These arguments can be self-serving, and specious too. Nevertheless, for many years the *modus operandi* worked acceptably. Deposit insurance premiums were low

14. For a less-than-flattering portrait of the Federal Reserve, see Greider (1987).

15. Some [see Golembe (1960)] prefer to describe the program as a government guarantee rather than insurance since the program is ultimately backed by the government's power to tax rather than by a finite pool of resources contributed by the insured.

16. Coverage of collectively owned accounts (for example, the deposit of a pension fund) can be far greater than $100,000. These accounts are protected for $100,000 multiplied by the number of participants in the collective.

(6 to 12 basis points) per annum, levied against insured as well as uninsured domestic deposits. Bank failures were few (averaging less than 10 per year before 1975, and almost all of these were small banks), and the deposit insurance funds (the FDIC's and the FSLIC's) grew steadily, until the 1980s. The FDIC was a regulatory backwater during most of this time, rarely seen or heard from. Its anonymity was testimony to its success.

This changed when bank failures ran at 200 per year in the late 1980s, with very large banks represented among the failed. These traumas led to insurance premiums rising to about 25 basis points in 1993. The visibility of the FDIC grew with adversity. The FDIC Bank Insurance Fund (BIF) was in deficit in 1991 and 1992, but recovered in 1993, and since then has grown steadily. See Figures 11.2 and 11.3.

The FDIC now works on a measure of the FDIC's fund balance as a percentage of the total deposits insured. It is called the *reserve ratio*, defined as:

$$\text{Reserve Ratio for a year} = \frac{\text{Fourth-Quarter Fund Balance}}{\text{Fourth-Quarter Estimated Insured Deposits}}$$

Figure 11.4 shows the behavior of the reserve ratio through time. It stood at 1.25 percent as of December 2005. Regulation requires that the designated (minimum) reserve ratio is 1.25 percent. The deposit insurance assessment rate during 2006 ranged between 0 and 0.27%, with most banks paying no deposit insurance premiums and the average annual assessment at roughly 0.11%. United States banking law requires a premium of 0.23% if the reserve ratio falls below 1.25.

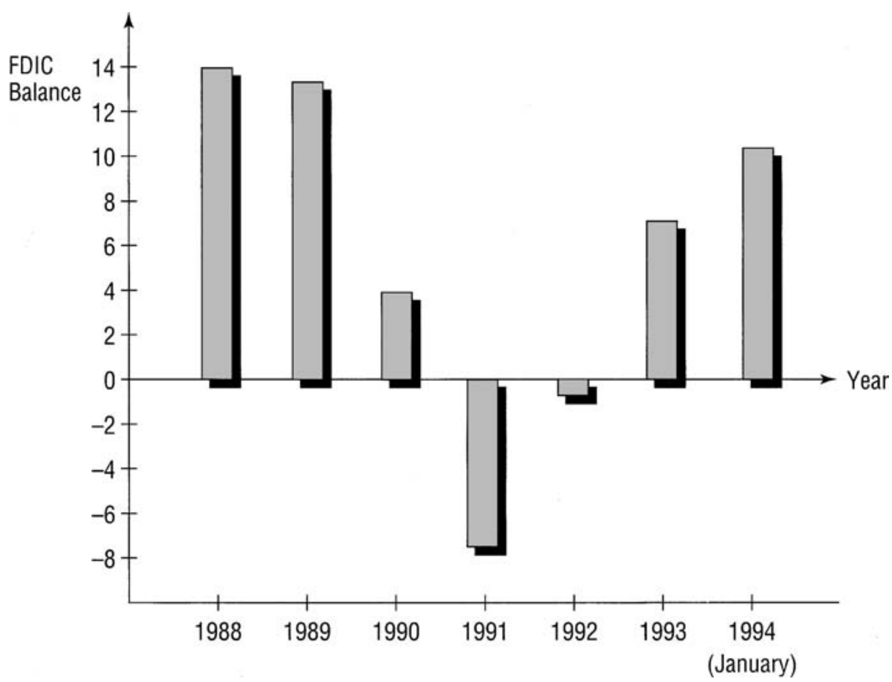


FIGURE 11.2 FDIC Bank Insurance Fund, in Billions of Dollars, During Resolution of Banking Gaisis and Just After
Source: FDIC.

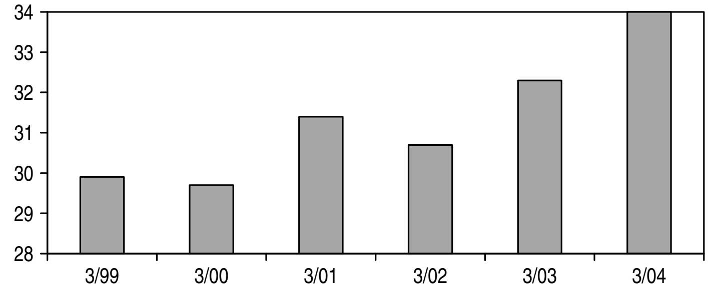


FIGURE 11.3 Bank Insurance Fund During 1999–2004
Source: FDIC.

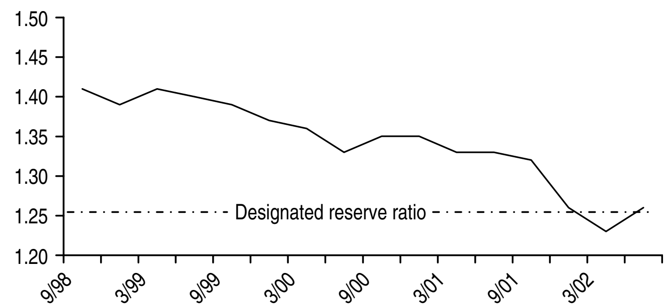


FIGURE 11.4 The Bank Insurance Fund Reserve Ratio Through Time
Source: FDIC.

Other Federal Agencies That Regulate Banks

Nonbank government agencies involved in banking include the Environmental Protection Agency, the Department of Labor, the Internal Revenue Service, and the Federal Bureau of Investigation (every cash transaction of $10,000 or more, with certain exceptions, must be reported to the FBI). However, these agencies are less involved in the day-to-day operations of banks than the banking agencies.

Noteworthy, however, are the Securities and Exchange Commission and the Antitrust Division of the Department of Justice. The former regulates the sale of debt and equity securities in public markets. The latter has responsibility for administering antitrust laws in cases of bank mergers and acquisitions, and in cases of collusion and other anticompetitive behaviors.

The SEC has asserted itself recently in the debate over bank accounting, proposing the replacement of GAAP accounting with market value accounting. This is an SEC issue because it relates to financial reporting and disclosure. We shall return to this issue later.

Market Structure and Competition

This regulation addresses issues of anticompetitive behavior among depository institutions. However, there is a *tension* between competition and safety regulation that relates to the problems in the banking and thrift industries in the 1990s.

Regulation impedes competition by elevating barriers to entry, branching restrictions, bank holding company limitations, and merger controls. You will recall from Chapter 10 that safety is enhanced by improving the charter values of existing banks and thrifts. This calls for *limiting entry*, so as to increase the economic rents earned by incumbents. Both public policy and technological advances have lowered entry barriers in recent years. In recent years, the tension between competition and safety regulation was resolved in favor of the former, with predictable consequences. We now turn to the ways in which market structure and competition regulation can affect industry structure.

- **Bank Chartering:** Chartering policy is designed to influence industry structure and also to foster adequate capitalization as well as ethical and competent management. Prior to the FDIC, entry controls were shared by the OCC and the state banking agencies and consequently varied widely. The advent of the FDIC added a measure of uniformity to the standards for chartering banks and thrifts. The FDIC, the Fed, and the OCC collaborated to implement a restrictive chartering policy until the mid-1960s.[17] For example, only 70 new banks were chartered between 1936 and 1955. More recently, however, chartering requirements have been eased, and now a well-designed operating plan along with adequate capital resources and credible management are usually sufficient to obtain a bank charter.
- **Branching and Bank Holding Company Restrictions:** The United States had over 11,000 banks in 1993, surpassing any other industrialized nation. The main reason for this proliferation of banks was a highly restrictive branching policy that existed until the Riegel-Neal Interstate Banking and Branching Efficiency Act of 1994 permitted unimpeded interstate banking. Both national- and state-chartered banks were limited in their geographic expansion by laws of their domicile states. This is the heritage of the McFadden Act of 1927. Pursuant to the U.S. Treasury's proposal of February 1991 (see Chapter 12), legislation was drafted to permit adequately capitalized banks to branch without regard to state boundaries, but the bill failed. To be sure, along with multistate pacts, many failed institutions provided opportunities for interstate expansion.[18] Thus, even prior to the Riegel-Neal Act, Citicorp claimed to do business in 30 states and Norwest (now Wells Fargo) in all 50 states through a patchwork of failed thrifts and banks they had

17. Besonko and Thakor (1992) examine the allocational consequences of a restrictive chartering policy.
18. Falling institutions provide an opportunity to circumvent state laws because the federal government can arrange sales of impaired institutions without regard to state restrictions.

purchased, along with mortgage banking and consumer finance companies. Others like Bank of America also built formidable interstate organization, but the pattern was checkered.

Before the Bank Holding Company Act of 1956, holding companies were permitted in a number of upper Midwest, West Coast, and Southern states and became an instrument for circumventing restrictions on branch banking. Thus, if a bank desired to expand in a state that severely restricted branching, it could establish a holding company that could, in turn, purchase separately incorporated banks within the state. Such a structure permitted the exploitation of economies of scale in marketing, finance, and processing, but each bank also needed to sustain the cost of being a separate corporate entity. Reserve requirements and capital requirements needed to be maintained separately, and each bank needed a separate board of directors. Clearly, the multibank holding company could not achieve all the potential savings of a branch structure.

The Bank Holding Company Act of 1956 brought the multibank holding company under the supervision of the Federal Reserve. All holding company formations and their bank acquisitions thereafter required the explicit permission of the Federal Reserve. What the 1956 legislation did not anticipate was the use of the bank holding company for purposes other than the purchase of banks. In the 1960s, there emerged a new kind of bank holding company, referred to as the one-bank holding company. It was used for a variety of circumventing purposes. For example, one-bank holding companies issued commercial paper, which banks were not permitted to do. They downstreamed the proceeds to their affiliate bank, which issued nondeposit liabilities for the proceeds of the commercial paper. The banks thereby avoided cash-asset reserve requirements against these liabilities.

The one-bank holding company also was a vehicle for increasing financial leverage. By purchasing the banks' equity with a mix of holding company debt and equity, the banks' owners increased their leverage. The holding company also was used as a tax shield in that dividends from the bank to the holding company could be used to retire holding company debt without being taxed as income to the holding company owners.

Finally, the one-bank holding company was used to expand the powers of banks. Bank holding companies purchased travel agencies, consulting companies, securities affiliates, and other businesses that banks would not have been permitted to purchase directly. Not surprisingly, one-bank holding companies experienced rapid growth after the 1956 legislation.[19] The 1970 Douglas Amendments to the Bank Holding Company Act brought one-bank holding companies under the supervision of the Federal Reserve. Thereafter, all holding company formations and all acquisitions, bank or nonbank entities, would require the explicit permission of the Federal Reserve. As indicated earlier, the vagueness of bank holding company legislation gave the Federal Reserve expansive discretionary powers. Prior to the bank capital requirements legislated in 1978, bank holding company applications became the Federal Reserve's foremost lever for coercing additional capital into the banks.

Bank holding company legislation of 1956 and 1970 clearly established the Federal Reserve as the primary federal bank regulatory agency.

19. See Fischer (1986).

Bank Holding Companies and Separability

Both public regulators and the banks themselves often prefer to lodge less traditional activities in a separately incorporated holding company subsidiary instead of having the bank itself engage in new businesses. The rationale is based on two considerations. First, the holding company is viewed as a "source of strength."[1] According to Regulation Y of the Federal Reserve, "A bank holding company shall serve as a source of financial and managerial strength to its subsidiary . . . " Second, prohibiting the bank from engaging directly in an activity achieves a measure of separation so that if something goes wrong at the new business, the bank, where most of the assets and net worth usually reside, will be insulated from the adversity.

But is it possible to insulate the bank in this fashion? This will depend on a variety of considerations. First, is the question of whether creditors of the subsidiary have legal remedy against the bank and/or the holding company. This is an issue the lawyers call "piercing the corporate veil." The courts usually respect the legal partitioning of related companies, but this depends on how the courts may act on such representations. The company's advertising may well influence the courts in deciding whether to respect the format separation. Thus, if a bank gives the public to understand that it stands behind the commitments of a subsidiary, the courts might feel justified in permitting creditors of the subsidiary to seek satisfaction from the bank or the holding company. The standards in this area are of necessity judgmental and less than clearly defined.

Perhaps even more important than legal considerations are the reputational issues. For example, will adversity at some nonbank subsidiary result in higher costs or lost business to other holding company affiliates? The failure of a subsidiary might lead to downgrading in the credit rating of the parent or the bank affiliate. In order to forestall such a possibility, the management of the parent might voluntarily divert resources to support the floundering subsidiary. Either one of these possibilities, the customer's reaction or the voluntary diversion of resources to support the floundering subsidiary, would subvert the separation achieved by the holding company structure.

Hence, the holding company's ability to insulate members can easily be overstated, and often is. Those who argue that banks can be permitted to do any legal businesses with impunity so long as the nonbank activities are isolated in holding company subsidiaries fail to appreciate the fragility of the separation provided by the holding company structure.

1. See Mester (1992).

Prior to the dismantling of the Glass-Steagall Act, the expanding securities activities of banks were being forced into holding company affiliates in order to achieve a measure of separation (see the box on Holding Companies and Separability) between the deposit-taking bank and its nontraditional activities.

The passage of the Reigel-Neal Act in 1994 finally permitted banks to branch across state lines. However, given the overcapacity in bank branches at that time, the way that banks expanded subsequently across state lines was by acquiring banks in various states rather than opening new branches. An important consequence of this consolidation was a dramatic decline in the number of U.S. banks, which stood at 7,435 by September 30, 2005. See Figure 11.5.

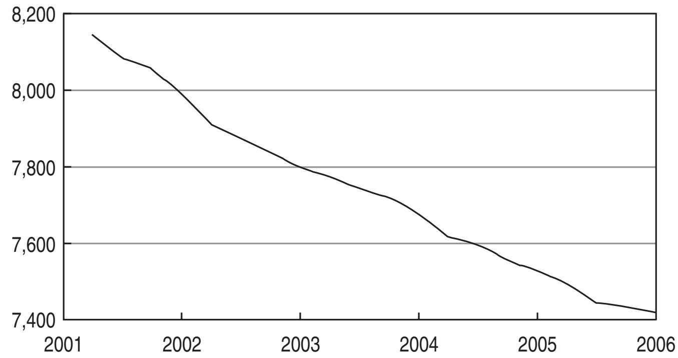


FIGURE 11.5 Number of Commercial Banks in the U.S.
Source: Federal Financial Institutions Examination Council; 2006 Federal Reserve Bank of St. Louis, research.stlouisfed.org.

Merging with another bank is an alternative method of branching, but the Bank Merger Act of 1960 requires banking authorities to review all proposed mergers, and after obtaining the opinion of the Department of Justice regarding the anticompetitive effects. A 1966 amendment to the Bank Merger Act shifted more of the responsibility to the Department of Justice where anticompetitive considerations were elevated. Although guidelines were revised again in 1982, the basic idea is to determine whether a bank merger would significantly reduce competition.[20]

Safety and Soundness Regulation

Bank regulation seeks to promote a "safe and sound" banking system. Regulators have been severely challenged on this dimension in recent decades. We will discuss the possible reasons for failures of safety regulation in the next chapter. Here, we address safety regulation as it is *supposed* to work. Figure 11.6 summarizes how safety regulations support bank safety. Redundancy in these regulations recognizes the difficulty of achieving safety objectives, especially when the regulated institutions can circumvent these regulations.

- **Federal Deposit Insurance:** In Chapter 10 we explain how deposit insurance creates a risk-inducing moral hazard. Thus, restrictive regulations are deployed to mitigate these endogenous risks.

20. In measuring competition, the Department of Justice includes thrifts if they are engaged primarily in retail banking. The courts viewed the Bank Merger Act and amendments as applying the Clayton and Sherman acts standards to banks.

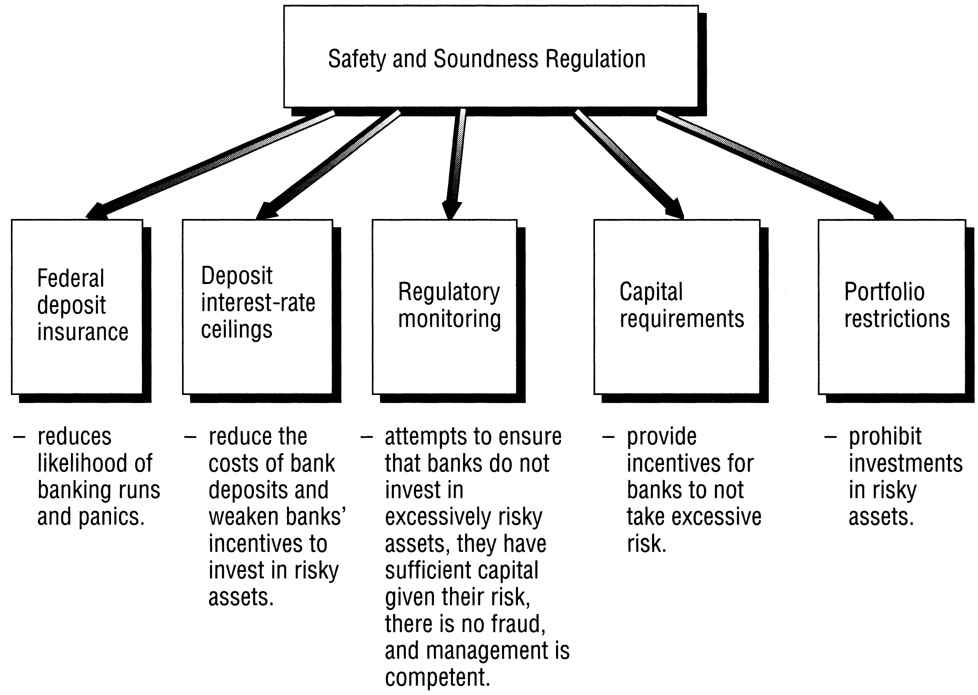


FIGURE 11.6 Types of Safety and Soundness Regulation

- **Deposit Interest-Rate Ceilings:** Deposit interest-rate ceilings, enforced via Regulation Q, came into existence with the Banking Act of 1933. Payment of interest on demand deposits was prohibited and the Federal Reserve was authorized to impose ceilings on the interest rates paid on time and savings deposits by member banks. *Regulation Q* ("Reg Q" for short) was subsequently extended to all FDIC-insured banks.

  The Reg Q ceilings lowered deposit costs and thereby offset unpopular deposit insurance premiums. Deposit interest-rate ceilings were defended on two grounds. First, it was thought that the incentives of banks to invest in risky assets would be weakened, and bank's profits would be both higher and less volatile. Second, it was believed that if interest rates were not restricted, larger banks in the money centers would attract deposits away from rural areas.

  Whether deposit interest-rate ceilings can achieve either of these objectives depends on the effectiveness of *nonprice* competition. When a bank wishes to attract more deposits and finds the Reg Q ceilings binding, it will complete on other dimensions. During periods in which Reg Q ceilings were binding (such as 1969–70, 1973–74, and 1978–80), banks engaged in nonprice competition, ranging from merchandise giveaways to subsidized cash management services.[21]

  Of course, the fact that banks and other depository institutions could circumvent deposit interest-rate ceilings does not mean that the ceilings were without effect. The nonprice competition induced by the ceilings distorted the allocation of resources.

21. Other manifestations of nonprice competition included oversupply of branches, automated teller machines (ATMs), and additional hours during which banks were kept open.

To see this, note that banks could have offered nonprice inducements to depositors even in the absence of ceilings, and it is possible that the *unconstrained* optimal allocation involves both the payment of explicit interest and the provision of other services to depositors. But when the Reg Q ceiling is binding, *more* resources will be allocated to the provision of these ancillary services, and this additional allocation represents an almost certainly inefficient distortion. Moreover, like reserve requirements, Reg Q ceilings induced innovation of new liability instruments. Eurodollar deposit growth was stimulated by both reserve requirements and Reg Q.

By 1986, virtually all deposit interest ceilings were phased out.[22] Massive deposit outflows (disintermediation) due to the increasing disparity between market interest rates and the Reg Q ceilings were one justification for the phaseout. Nevertheless, the unfettered competition for deposits following the removal of Reg Q contributed to a decline in bank profitability.[23]

- **Regulatory Monitoring:** The periodic examination of banks by the public regulatory agencies is a central part of regulation. Indeed, each of the three federal bank regulatory agencies employs bank examiners and each of the state banking agencies has theirs as well.

Although bank examiners have overlapping jurisdictions, there is a formal division of labor. The OCC examines all national banks, the Federal Reserve all state-chartered Federal Reserve member banks, and the FDIC all remaining insured banks. Of course, the states are responsible for examining all state-chartered banks, too. Often state and federal agencies accept each other's exams. Sometimes they examine jointly. There is some coordination, along with considerable redundancy. Examination details are provided in the box below.

Regulatory Rating System

A uniform interagency bank rating system known as CAMEL (capital adequacy (C), asset quality (A), management ability (M), earning quality (E), liquidity level (L)) was adopted in 1978. In 1997, a sixth factor, "Sensitivity to Market Risk" (S), was added, to make it CAMELS. None of these factors is judged in isolation. For example, what is acceptable asset quality will depend on how much capital the bank has.

- **Capital Adequacy:** The bank's capital is evaluated on the basis of both the bank's sizes as well as the composition of its assets and liabilities, both on- and off-balance sheet. We will have more to say about capital shortly.

22. Banking institutions are still prohibited from paying any interest on demand deposits. However, this rule is of little consequence because consumer transaction accounts are classified as NOW or share draft accounts. Since NOW accounts are available to individuals, nonprofit entities, and public agencies, the prohibition on demand deposit interest payments applies only to business deposits. Banks may, however, provide their business customers with cash management, lockbox, payroll, and similar services without violating the interest payment prohibition. See Huber (1989).

23. It has been estimated that banks gave back about half of their savings from deposit ceilings in the form of services. See Flannery (1983).

- **Asset Quality:** Examiners assess the credit risks in the various loans in the bank's portfolio and classify these loans as: good, substandard, doubtful, or loss.
- **Management Ability:** Examiners attempt to gauge not only the bank's management but also its board of directors. Competence, management acumen, integrity, and willingness to comply with banking regulations are some of the factors assessed.
- **Earnings:** There is an evaluation of the earnings as well as their level relative to peers. One objective is to assess the impact on the bank's capital of internally generated funds.
- **Liquidity:** Regulators assess liquidity by examining credit conditions, deposit volatility, loan commitments, and other contingent claims against the bank, capital, current stock of liquid assets, and the bank's perceived ability to raise funds on short notice.
- **Sensitivity to Market Risk:** Regulators assess how sensitive the bank's asset, liability and net worth values are to changes in market condition like interest rates.

Until recently, regulatory examinations served in lieu of external audits for most banks. However, bank examinations are not the equivalent of an external audit. External audits focus on financial reporting and consistency with GAAP and are put to external as well as internal use. Bank examinations focus on asset quality and the effectiveness of monitoring and are for internal use exclusively. Indeed, it is illegal to disclose bank examinations (for example, CAMELS ratings) outside the bank. Both regulatory exams and external audits also seek to uncover fraud, but some would say with checked success.[24]

The Federal Deposit Insurance Corporation Improvement Act (FDICIA) of 1991 mandates annual, full-scope examinations of banks by regulators. These regulatory examinations and audits are predicated on the assumed informativeness of ratings-based classifications resulting from the examinations in terms of their ability to reveal the bank's true financial health. How informative CAMELS ratings are in assessing a bank's financial condition is an open question. It seems plausible that during stable periods in the banking industry, these ratings are more informative about a bank's financial conditions than during times of stress.[25]

The result of the bank's regulatory examination are reported to the bank's board of directors, with subsequent dialogue between the examiners and directors to clarify issues and to discuss steps for dealing with the problems uncovered by the examination. The examination report is then submitted to the supervisory authorities whose relationship with the bank is guided by the findings of the report.

Supervisors can impose wide-ranging sanctions for improper actions by the management or the board. Advice is followed by warnings, then cease and desist orders. Management can be discharged. Directors can be fined, discharged, and

24. Banks and bank holding companies have been required to have audits if shareholders exceed 500, but these are SEC requirements. Only with the 1991 Federal Deposit Insurance Corporation Improvement Act (FDICIA) were banks (holding companies) with assets exceeding $150 million required to have external audits.

25. This is precisely what Gasbarro, Sadguna and Zumwalt (2002) found for Indonesian banks during the Southeast Asian financial crisis. While most of the CAMELS variables were informative during periods of stability, only the earnings variable mattered during a crisis.

barred from banking. The bank can be put into conservatorship or into receivership. Directors and officers can be sued civilly and/or criminally for failure to discharge fiduciary responsibilities, for negligence, gross negligence, or criminal negligence. The regulators' power to impose sanctions is expansive.

How informative CAMELS ratings are in assessing a bank's financial condition is an open question. It seems plausible that during stable periods in the banking industry, these ratings are more informative about banks' financial conditions than during times of crises when some of the factors in CAMELS may be less revelant.[26]

- **Capital Requirements:** In book value terms, capital is the sum of retained earnings and the purchase price of outstanding common stock, whereas in market value terms it is the current market price per share multiplied by the total number of shares outstanding. For regulatory purposes, however, capital is defined also to include general, but not specific, *loan loss reserves*, permanent preferred stock, and certain *long-term debt*.[27] Loan loss reserves are capital that has been earmarked to absorb future loan losses; when these losses occur, they are charged against the loan loss reserve account rather than against current earnings. This practice of reserving for losses smoothes the time pattern of income, a practice that has become unpopular in recent years. As a form of capital, loan loss reserves improve the value of the bank's creditors' claims. Similarly, long-term debt (which includes mainly subordinated notes and debentures) is junior to deposits, so that a greater amount of long-term debt on the bank's balance sheet implies greater protection for depositors. Hence, the regulatory rationale for including loan loss reserves and long-term debt in bank capital is that both are junior to deposit therefore serve to protect the depositor as well as the deposit insurer. Under recently adopted rules, federal bank regulators make a distinction between Tier-1 and Tier-2 capital. Capital requirements apply to total (Tier-1 plus Tier-2) capital, with an upper limit on the amount of Tier-2 capital contributing to the total. In the box below, these two types of capital are defined.

Tier-1 and Tier-2 Capital

Tier-1 (Core) Capital:
- Common Stock
- Retained Earnings
- Capital Surplus (amount received from sale of common stock or preferred stock in excess of par)
- Disclosed Capital Reserves (reserves set aside for cash dividends not declared plus amounts for unforeseen contingencies)

26. This is precisely what Gasbarro, Sadguna, and Zumwalt (2002) found for Indonesian banks during the Southeast Asian financial crisis. While most of the CAMELS variables were informative during Indonesia's stable economic periods, the informativeness of the CAMELS variables declines during crisis periods.

27. Specific reserves are dedicated to a particular impaired asset, whereas general reserves are not assigned to identifiably impaired assets, and are therefore generally available.

Tier-2 (Supplementary) Capital:
- Loan and Lease Loss Allowances
- Preferred Stock with Maturity of at Least 20 Years
- Subordinated Obligations (Both Stock and Debt) With an Original Average Maturity of at Least 7 Years
- Undisclosed Capital Reserves
- Hybrid Capital Instruments

Total Capital:
Tier-1 Capital + Tier-2 Capital

Until the 1980s, legal cash-asset reserve requirements were a more important constraint on bank's balance sheets than capital requirements. To be sure, newly chartered banks had reasonably well-defined initial capital requirements, and the Federal Reserve Bank of New York had a capital standard that was similar to the present risk-related capital rules. Nevertheless, bank capital regulation was in striking contrast to thrift regulation where capital requirements had primacy over liquidity requirements as a regulatory desideratum.

This contrast reflected the traditional view that the principal risk in banking was withdrawal risk rather than credit risk, and the reverse was the presumption regarding thrifts. Commercial banks were designed to make short-term, self-liquidating business loans, financing trade, inventories, and receivables with highly predictable patterns of repayment. But LBO, LDC, commercial real estate, term lending, off-balance sheet activities, and bank holding company extensions into nonbank activities expanded the credit risk asset transformation of commercial banks.

The banks' expanded credit risk implied greater exposure of the governmental safety net, and the regulatory response was more stringent bank capital requirements. Bank shareholders were required to put up greater stakes in order to control banking assets.

The first effective nationwide capital requirement for commercial banks was mandated by the International Banking Act (IBA) of 1978. Previously, the only federal capital standards were for newly chartered banks. Otherwise, capital standards were *ad hoc*, usually implemented as an incident to a BHC application. For example, if a BHC sought permission to acquire a mortgage banking affiliate, the Federal Reserve would require additional capital in order to enter this new line of business.

The 1978 legislation required bank capital of at least 5.5 percent of total assets, and capital was defined to include paid-in equity, retained earnings, general (but not specific) loan-loss reserves, limited amounts of permanent stock, and certain classes of convertible long-term debt.

The Basel I Capital Accord

Since 1978, bank capital has become a focal point of bank regulation. With increasing international competition among banks in a global market, public regulators have come to recognize the need to coordinate capital requirements for banks across countries. Hence, meetings were held among the United States, Japan, and the major Western European countries under the auspices of the *Bank for International*

TABLE 11.1 The Basel I (BIS) Capital Requirements

Minimum Overall Capital Ratio: 8 percent
Mix in Capital Ratio: Not more than 50 percent Tier-2 Capital
Requirements Against Specific Assets

Asset Risk Category	Conversion Factor in Percentage	Qualifying Assets
1	0	Cash (including foreign currency), claims on Federal Reserve Banks, direct obligations of the U.S. with a maturity of up to 91 days, claims on OECD central government and banks, and loan commitments with maturities less than 1 year.
2	10	Longer-term federal government debt, loans secured by government paper or deposits at the official lending institution, and Federal Reserve System bank stock (at book value).
3	20	Claims on domestic depository institutions, short-term claims on foreign banks in OECD countries, cash items in the collection process, obligations of or claims guaranteed by federal entities, claims backed by the full faith and credit of state and local governments, and the lowest-risk standby letters of credit.
4	50	Government obligations whose repayment is not backed by the full faith and credit of the issuing entity (revenue bonds and similar paper), residential mortgages, unused loan commitments with maturities exceeding 1 year, note issuance facilities, and medium-risk standby letters of credit.
5	100	Claims on corporations (including loans and bonds), guaranty-type instruments, sales subject to repurchase agreements and other credit substitutes, and certain standby letters of credit.

Settlements in Basel, Switzerland, in 1987. After long and arduous negotiations, the Basel (BIS) Accord provided a more-or-less uniform capital standard for all banks in the 12 participant countries. The new BIS (or Cooke, after the name of the British organizer) ratios were to be fully implemented by 1993 and to cover all insured banks. The accord, now referred to as the Basel I Accord, was lauded as a great victory in international banking cooperation and as the harbinger of the "level playing field." Its main elements are summarized in Table 11.1. The definitions of Tier-1 capital are those given in the box earlier. The guidelines specifically exclude the following items (included in earlier capital ratio calculations) from the capital base for computing capital ratios: (i) goodwill, (ii) other intangibles, (iii) capital investments in most unconsolidated subsidiaries, (iv) reciprocal holdings of capital instruments in banking organizations, and (v) revaluation reserves.

The Basel I Accord relates required capital to the composition of the bank's assets. Hence, capital requirements are stated as a percentage of *risk-weighted assets* rather than total assets. A bank's risk-weighted assets are an average of the bank's booked assets and *credit equivalent amounts* of its off-balance sheet exposure. There are five asset categories for risk-weighting purposes, numbered 1 through 5 in Table 11.1. The Basel I Accord called for a minimum overall risk-weighted capital ratio of 8 percent, with at least 50 percent in the form of Tier-1 capital. To compute how much capital it needs, a bank must first determine the dollar volume of assets in each of the five risk categories. Say $\$A_i$ represents the dollar volume in risk category i, with $i = 1,2,3,4$, or 5. Let C_i represent the conversion factor for category i, that is, $C_1 = 0$, $C_2 = 0.10$, $C_3 = 0.2$, $C_4 = 0.5$, and $C_5 = 1.0$. Then the total capital a bank is required to have is 8 percent of $C_1A_1 + C_2A_2 + C_3A_3 + C_4A_4 + C_5A_5$ or $0.08 \times (C_2A_2 +$

TABLE 11.2 Summary of Capital Constraints on U.S. Banks

1. Total capital $\geq 0.008 \times \sum_{i=2}^{5} C_i A_i$
2. Tier-1 capital $\geq 0.04 \times \sum_{i=2}^{5} C_i A_i$
3. Tier-1 capital $\geq \left[0.03 \times \sum_{i=2}^{5} A_i\right] + d$

where A_i = dollar volume of assets in category i, and i goes from 1 to 5 as in Table 11.1.

C_i = Basel risk weight or conversion factor attached to category i.

d = an add-on usually between $0.02\Sigma A_i$ and $0.03\Sigma A_i$ for banks with CAMEL ratings below the best

Note that (2) and (3) can be combined as:

4. Tier-1 capital $\geq \text{Max}\left\{0.04 \times \sum_{i=2}^{5} C_i A_i, \left[\left(0.03 \times \sum_{i=2}^{5} A_i\right)\right] + d\right]\right\}$

where "Max" means the greater of the two enclosed quantities in the parenthesis. The bank's capital must then satisfy both (1) and (4).

$C_3A_3 + C_4A_4 + C_5A_5$) since $C_1 = 0$. Moreover, the bank must have at least $0.04 \times (C_2A_2 + C_3A_3 + C_4A_4 + C_5A_5)$ as Tier-1 capital. U.S. banking regulators have additionally imposed a Prompt Corrective Action (PCA) requirement in terms of a leverage ratio constraint mandating that Tier-1 capital can be no lower than 3 percent of *total assets* for banks earning the highest CAMEL rating. Other banks are required to keep a 4 percent ratio. Weaker-rated banks may have to keep higher capital, as much as 6 percent. Table 11.2 summarizes the three capital constraints U.S. banks face.

Signatories to the Basel I Accord were free to impose higher capital requirements on banks in their own countries. As indicated previously, the capital requirements are risk sensitive for various classes of assets, both on- and off-balance sheet items. Although the focus of the Basel I capital requirement is credit risk, limited recognition is also made of interest-rate risk. For example, the capital requirement on federal government debt with initial maturity exceeding 91 days is 0.8 percent, whereas the requirement on shorter-maturity government debts is 0. The imposition of capital requirements on shorter-maturity government debts is 0. The imposition of capital requirements against off-balance sheet items (for example, loan commitments, standby L/Cs, and interest-rate and currency swaps) is another innovation of the Basel I requirements.

United States banks can be fined $1,000 per day for violating these capital requirements. Moreover, the OTS, which administers capital guidelines for insured thrifts, has adopted analogous risk-based capital requirements.[28]

By 1993, all of the world's major banks had satisfied the Basel capital requirements.[29] Large American banks had been increasing their capital ratios ("deleveraging") since 1989. See Figure 11.7.

From one perspective, capital requirements look like a naïve first step in international banking cooperation. There are numerous criticisms. First, the risk classes

28. Thrifts are also subject to an 8 percent capital requirement. Core capital cannot be lower than 3 percent, but this minimum applies only to the best capitalized thrifts. Others will have a 4 percent core capital requirement. The risk-weighting of assets for thrifts is different from that for banks. Cash and U.S. government obligations have zero weight, mortgage-related obligations guaranteed by the United States, and obligations backed by the full faith and credit of state and local governments have a 20 percent risk weight. State revenue obligations and residential mortgages have a 50 percent risk weight, other mortgages have a 100 percent risk weight, goodwill and real estate owned by a thrift have a 200 percent risk weight, and equity investments have a 300 percent risk weight.

29. See Reuters (1993).

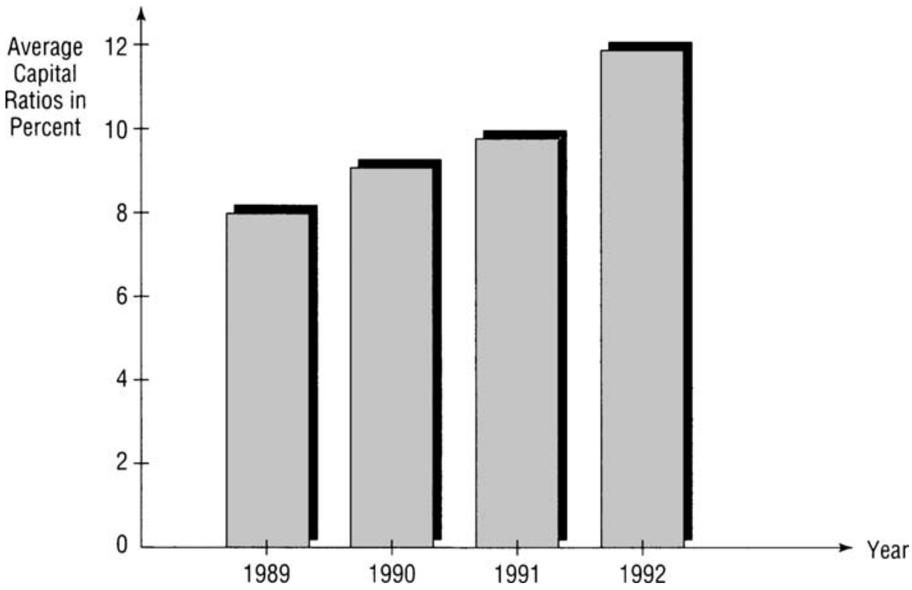


FIGURE 11.7 Average Capital Ratios for Large United States Banks
Source: The Economist (1993).

are crude to the point of inviting exploitation. Mortgages require half the capital of business loans, yet it is easy to find mortgages with greater credit risk than business loans. Indeed, at the margin, business loans can be repackaged in the form of mortgages. Regulators always seem to underestimate the adroitness and plasticity of the capital markets. Loans are merely written contracts that can be adapted to meet the imperatives of the moment.

Second, the risk classes can be manipulated.[30] For instance, suppose a bank invests in U.S. Treasury bonds that require zero capital and then enters into an amortizing swap (recall Chapter 8) in which it *pays* the total return on those bonds and *receives* the total payments on mortgages. Even though this bank effectively holds mortgages, the bank faces a BIS capital requirement that is lower than the 4 percent attached to mortgages.

Third, interest-rate risk failed to receive its due under Basel I, although the 1991 legislation mandates that regulators develop new capital guidelines that reflect interest-rate risk as well as credit risk.

Fourth, many concessions were made to accommodate special interests. For example, 45 percent of unrecorded capital gains on equity holdings can be counted as Tier-2 capital. This concession was especially important to Japanese banks with large equity holdings valued at purchase prices. U.S. banks are prohibited from treating unrealized capital gains as capital.

Fifth, the Basel I capital requirements assume that banking risk is substantially the same in different countries. However, there are striking differences in the variability of bank rates of return across countries. This suggests that the basic asset-risk categories may be too crude and that minimum capital ratios should vary across countries.

Finally, since the capital ratios are prescribed on a book-value basis, they fail to adjust for changing return volatilities and the relationship between book and market

30. This was pointed out by Merton (1994).

values of bank equity. Moreover, the capital requirements do not recognize the portfolio aspects of bank balance sheets. Since requirements are linear in individual asset categories, there is no recognition of the covariability of returns that affects diversification and portfolio risk.

Despite these shortcomings, the accord is noteworthy as a first step in the international harmonization of capital standards, linking capital to risk, and in recognizing the significance of off-balance sheet items. As it turned out, it was basically a first step, as a revised accord, the Basel II Accord, was adopted in 2004; this will be discussed in the next chapter. Note also that the prompt corrective action requirements of FDICIA increases the importance of capital since regulators are required to close banks with sufficiently deficient capital. We turn now to the issue of *why* capital requirements may be important, and how an optimal capital standard might be determined.

- **The Effects of Capital Requirements:** Capital controls the bank's appetite for risk-taking through a *direct* as well as an *indirect effect*. As a buffer against loan and security losses, capital acts like a deductible in an insurance policy. In addition, capital can reduce a bank's incentive to invest in risky assets because the amount of debt in the bank's capital structure is being reduced. This indirect incentive effect of bank capital is not necessarily monotonic. Figure 11.8 depicts the likely nonmonotonic effect of capital on bank risk. A distinction between the bank's "insiders" (top bank managers who own some equity) and "other shareholders" (not involved in management) is assumed. The bank's decision making is likely to maximize the welfare of insiders, even at the expense of other shareholders.[31] This

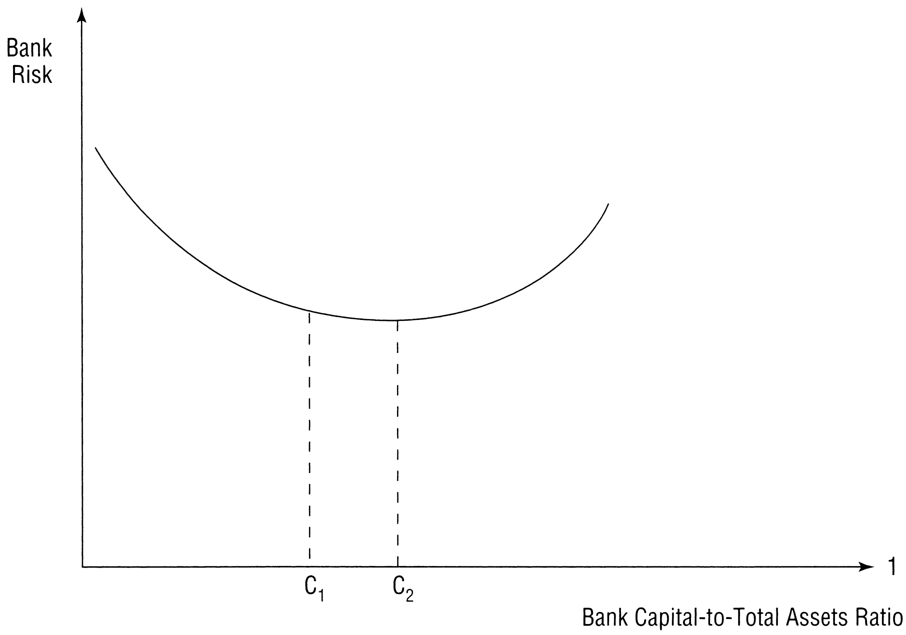


FIGURE 11.8 Possible Effect of Bank Capital on Bank Risk

31. This is the familiar agency problem between the manager and shareholders. Since managers own some stock, their incentives are somewhat aligned with those of the nonmanaging shareholders. However, moral hazard will remain as long as managers do not own *all* of the stock.

agency problem is the key to understanding Figure 11.8, which calls into question traditional thinking about the risk-deterring effects of bank capital.[32]

To see this, note that increasing the capital-to-total-assets ratio has *three* effects on the bank, two of which are the traditional buffer and incentive effects that lead to greater safety. The third arises from the agency problem between insiders and other shareholders. As the required capital-to-total-assets ratio rises, insiders may be unable to provide all of needed equity. Hence, increasing amounts of "outside" equity will be sought. As management's ownership position is diluted, their incentive to expend costly effort in managing the bank may diminish as well. The result may be less vigilant screening of loan applications, insufficient attention to risk management, and possibly increased consumption of perquisites. This could lead to an *increase* in bank risk as the capital ratio is raised beyond some point. We shall refer to this as the "outsider equity effect."

The summation of the buffer, incentive, and outside equity effects is depicted in Figure 11.8. The bank's risk is likely to decline with increases in the capital ratio up to some level, say C_1, as the buffer and incentive effects work to reduce risk and the outside equity effect has not yet come into play because insiders provide the necessary capital. Beyond C_1, however, outside equity is required. Between C_1 and C_2, even though the outside equity effect provides risk escalation incentives, the effect is still dominated by the buffer and incentive effects. But as the capital ratio rises above C_2, the outside equity effect eventually dominates the other two effects. We therefore get an increase in the bank's overall risk with increased capital.

The research on which this argument is based was conducted prior to the advent of risk-based capital requirements. Moreover, it applies narrowly to circumstances in which the bank is managed by an owner-manager with sufficient equity in the bank for the "ownership dilution effect" to be an important driver of the value of the bank. More recent research strongly suggests that a sufficiently high level of risk-based capital requirements can be an effective antidote to value-dissipating risk-shifting on the part of banks, and can enhance both the value of the individual bank as well as industry stability.[33]

Empirical Evidence on the Effects of the Basel I Accord and FDICIA of 1991

Ultimately, the effect of capital requirements should be assessed empirically. In this respect, the Basel I Accord, along with the FDICIA of 1991, represent a big success. *Figure 11.9* shows that bank capital ratios increased quite significantly after these regulatory initiatives.

Moreover, the increase in bank capital was accompanied by a dramatic decline in bank failures. See *Figure 11.10*. There were less than 10 commercial bank failures each year during 1995–2001. Moreover, the total level of protection against credit losses—defined as the percentage of assets represented by income, reserves and equity capital—has grown steadily through time. See *Figure 11.11*.

32. See Gennotte and Pyle (1991), Kahane (1977), and Koehn and Santomero (1980) for theoretical models that predict that higher capital requirements could increase bank risk. The argument presented here is based on Besanko and Kanatas (1991). The incentive effects of capital are also explored in Boot and Greenbaum (1992).

33. See Repullo (2004) for an excellent treatment of this issue.

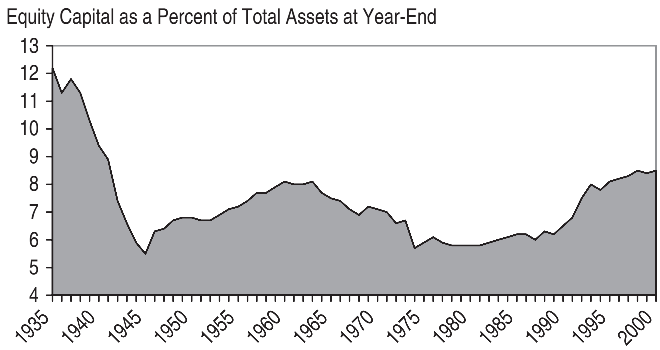


FIGURE 11.9 U.S. Bank Capital Ratios
Source: FDIC Historical Statistics on Banking.

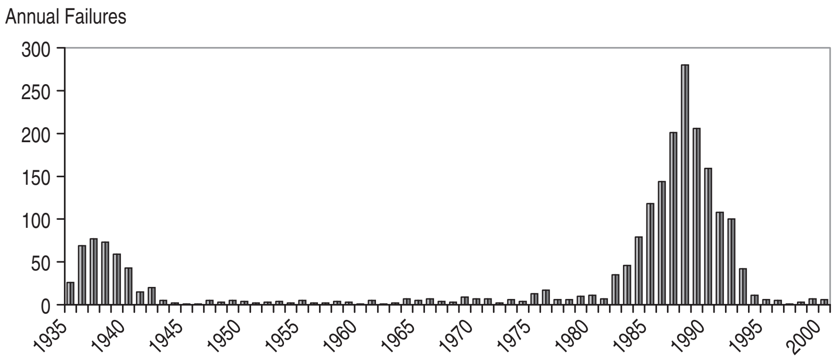


FIGURE 11.10 Annual Failures of FDIC-Insured Commercial Banks Remain Well Below Crisis Levels
Source: FDIC Historical Statistics on Banking.

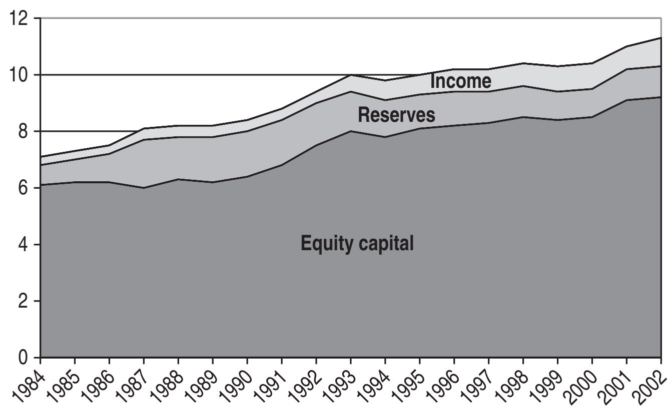


FIGURE 11.11 Total Level of Protection Against Credit Losses Has Grown Steadily Over Time (all FDIC-insured institutions)
Source: FDIC Bank Call Reports.

Interestingly, as banks have increased their capital ratios, they have also increased their loans and reduced their holdings of low-risk securities. See *Figure 11.12*. This can be seen as a reflection of greater perceived risk-absorption ability due to higher capital, i.e., a riskier asset portfolio mix can be seen as a complement to reduced leverage.

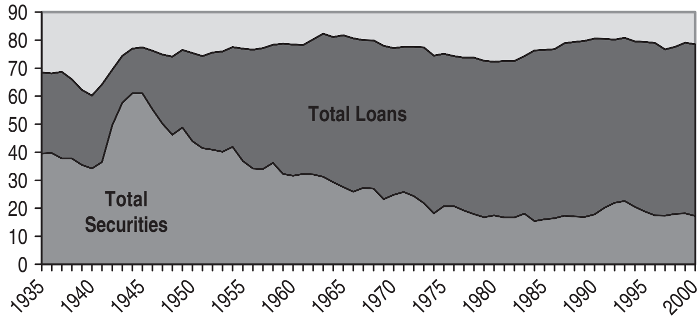


FIGURE 11.12 Bank Asset Mix Has Trended Toward Loans and Away from Low-Risk Securities (FDIC-insured commercial banks)
Source: FDIC Historical Statistics on Banking.

Safety and Soundness Regulation: Bank Portfolio Restrictions

The legal separation of commercial and investment banking in the United States was memorialized with the Glass-Steagall Act of 1933.[34] Federal Reserve member banks were prohibited from underwriting, distributing, or dealing in stocks, bonds, or other securities, the exceptions being U.S. government bonds, general obligation municipal securities, and the obligations of specified government agencies. The act also prohibited banks from affiliating with investment banking firms or otherwise engaging in investment banking. The separation of commercial and investment banking was based on the controversial notion that the massive banking disruption of the period was due to the securities activities of banks.

Glass-Steagall also affirmed the authority of the states regarding geographic expansion, originally spelled out by the McFadden Act in 1927. Thus, branching and holding company issues were deferred to the states, effectively blocking the development of interstate banking and ensuring a fragmented industry. The assets that banks could intermediate as well as their geographic origin were severely restricted. This meant that banks were closely tied to the fortunes of their local communities, and the opportunities for diversification were limited, as were opportunities to exploit economies of scale and scope.[35] Notably, the restrictions were based on an increasingly tenuous distinction between loans and securities. Private placements of debt securities are, for all practical purposes, loans, and securitized loans are, for all practical purposes, securities. The all-but-vacuous distinction between loans and debt securities was not lost on bankers seeking to expand their activities.

A relentless testing of limits by bankers prompted regulatory reinterpretations through time, leading to bank entry into a variety of previously prohibited areas. For example, through holding company affiliates, banks were able to underwrite municipal debt, commercial paper, and even corporate bonds and equity, within strict limit as to the volume of this business. Without the benefit of legislation, the rules govern-

34. See Benston and Kaufman (1988). Kahane (1977) was the first to show that neither capital requirements nor portfolio restrictions alone will be enough to control bank risk. Mester (1992) argues in favor of portfolio restrictions for somewhat different reasons.

35. As an offset to excessive concentration, banks are allowed, with minor exceptions, to lend to any one person (legal or biological) no more than 15 percent of their capital, but there are exceptions.

ing asset proscriptions were substantially relaxed by regulators. To be sure, these initiatives were tested in the courts, but the regulatory liberalizations were judicially sustained for the most part. This is actually quite remarkable because the reinterpretations of the 1930s legislation were fundamental. Underwriting corporate debt and equity securities by commercial banks and their holding companies was for decades simply illegal under Glass-Steagall. When the banks and their regulators sought to have the law liberalized, their efforts were frustrated time and again by a variety of other interest groups. Then, however, the bank regulators simply reinterpreted Glass-Steagall, and the courts upheld their prerogative to do so.

The once impregnable wall separating commercial and investment banking was dismantled piecemeal, without legislation. Likewise, banks found their way into the asset-management business via mutual funds. Banks could sell and manage mutual funds. This too was thought to be foreclosed by Glass-Steagall. The separation of banking and insurance proved to be more stubborn.[36] Lobbying by insurance interests kept banks out of this business for the most part. However, banks competed vigorously in the business of financial guarantees. Moreover, some insurance companies offered depository financial intermediation services.[37] Standby letters of credit sold by banks and financial guarantees sold by insurance companies are close substitutes, especially as credit enhancements for securitization.[38] Likewise, banks and insurance companies competed directly in the market for annuities.

Ultimately, the Glass-Steagall Act was dismantled in 1999, with the passage of the Financial Services Modernization Act of 1999. This act is also known as the Gramm-Leach-Bliley Act of 1999, to reflect the names of the senators who sponsored the bill. This legislation repealed Sections 20 and 32 of the Glass-Steagall Act. It further authorizes a wide range of activities for bank holding companies and foreign banks that meet eligibility criteria.

In the case of such organizations, it allows United States financial service providers, including banks, securities firms and insurance companies to affiliate with each other and enter into each other's markets. The affiliation of financial services providers allows open and free competition in the financial services industry. We will discuss this legislation in more detail in the next chapter.

Consumer Protection, Credit Allocation, and Monetary Control Regulation

Consumer Protection Regulation

Consumer protection regulation in the United States takes many forms. For example, *usury laws* restrict interest rates that lenders can charge on consumer loans. The idea of protecting borrowers from exploitative loan interest rates has biblical origins.

36. Banks are normally restricted to selling only credit-related insurance such as credit life, disability, and involuntary unemployment insurance, which may protect the bank's interest in the loans it grants. In towns with populations not greater than 5,000, however, banks can engage in other insurance activities if adequate insurance facilities are unavailable. Moreover, recent proposals would permit community development banks to offer insurance in low-income areas.

37. For example, some insurance companies own thrift institutions.

38. These comments apply to nationally chartered banks. Some states allow banks to engage in agency, brokerage, underwriting, and a broad range of insurance activities.

In the United States, legislation goes back to 1641 when Massachusetts passed a usury law. From an economic standpoint, usury ceilings do not make much sense unless the lender is a monopolist. When a bank encounters a borrower whose assessed risk warrants a loan interest rate higher than the usury ceiling, the bank will withhold credit.[39]

Another form of fairness regulation mandates the disclosure of information by lenders. The *Truth-in-Lending Act* (Federal Reserve Regulation Z) requires that lenders provide their customers with standardized credit information regarding finance charges and annual percentage rates in order to permit more informed borrower decisions. Similarly, the *Real Estate Settlement Procedures Act* (RESPA) requires that mortgage borrowers be provided all relevant information about the real estate settlement process, and a uniform settlement statement that discloses all fees and charges at closing.

In addition to protecting borrowers, legislation seeks to protect depositors. The *Truth in Savings Act* (TIS) was enacted in 1991 as part of the *Comprehensive Deposit Insurance Reform and Taxpayer Protection Act or Federal Deposit Insurance Corporation Improvement Act* (FDICIA) for short. The purpose of TIS, which went into effect June 1993, is to promote competition among depository decisions. TIS requires uniform disclosure of the terms and conditions for the payment of interest and the charging of deposit fees. It applies to all banks and thrifts, insured and uninsured. Credit unions are not directly subject to TIS, but the *National Credit Union Association* (NCUA)—the regulatory agency governing credit unions—is required to adopt similar rules.

Given the rapid growth in electronic funds transfer (EFT), it is not surprising that regulation governs this activity as well. The *Electronic Funds Transfer Act* (EFTA) of 1978 is designed to protect consumers by establishing the rights, liabilities, and responsibilities of EFT participants. The EFTA focuses on the types of transactions rather than the type of institution providing the service. It applies to most fund transfers initiated through an electronic terminal, telephone, or tape to authorize the debiting or crediting of an account by a financial institution. Thus, payment by check, for example, would not be covered by the EFTA.

The EFTA is a complex web of requirements, one of which is that the consumer be provided "means of access" (for example, a card and a personal identification number) that can be used to initiate an EFT transaction. Other requirements have to do with disclosure and documentation. Prior to a consumer's first EFT transaction, the financial institution must provide a written disclosure statement that clearly explains the terms and conditions under which the EFT service is provided. Moreover, for each transaction initiated at an electronic terminal, the financial institution must provide a written receipt that clearly states the relevant information about the transaction.

Credit Allocation Regulation

Because bank credit availability affects the pattern of economic activity, governments are often tempted to influence the allocation of this credit to achieve social and political objectives.[40] Some form of governmental credit allocation is found in almost

39. Robins (1974) conducted a study of real estate lending and found that when usury ceilings were lower than market mortgage rates, the level of residential construction declined by approximately 25 percent.

40. See Thakor and Beltz (1994).

every country. Atypically, *overt* governmental credit allocation has been used only sparingly in the United States, except during times of war or national emergencies. However, numerous indirect credit allocation mechanisms have been extensively employed in the United States. We discuss some of these below.

- **Credit for the Purchase of Securities:** Credit to finance securities has been regulated by the Federal Reserve since 1934. Initial credit is limited to a percentage of the value of the security. If the value of the security drops after credit is extended, a borrower is subject to a margin call. The borrower must then provide additional collateral or sell stock. Moreover, the Federal Reserve may limit a member bank's total lending in support of securities transactions. The obvious effect of this restriction and the margin requirements is to limit credit for the purchase of securities.[41]
- **Tax Policy and Guarantee Programs:** Tax credits and tax deductions have been used to influence a variety of economic activities, including credit allocation. For example, tax credits are used to encourage capital investment and greater energy efficiency in homes. Tax deductibility of charitable contributions has increased the flow of capital to eligible organizations.
- **Credit Programs for Specific Sectors of the Economy:** The government has a long history of credit programs to promote specific sectors of economy. Examples are housing, education, and agriculture.

  For example, the thrift industry was nurtured by the government primarily to encourage home ownership. The investment portfolios of these institutions were restricted and tax incentives were provided to encourage investments in residential mortgages. A major objective of Regulation Q was to keep funding costs low so that home buyers could obtain low-cost credit. Fannie Mae (FNMA), Ginnie Mae (GNMA), and Freddie Mac (FHLMC) were created and subsidized to provide a secondary market for mortgage loans in order to further encourage the flow of credit into housing. The government's role in directing credit to housing has been massive.

  In 1970, the federal government began directing credit to education with the adoption of the Guaranteed Student Loan Program. Loans were made available to students at favorable interest rates and liberal repayment terms. Sallie Mae (Student Loan Marketing association) was created to provide a secondary market for student loans.

  The federal government also subsidizes *agriculture*. The Farm Credit System (FCS) gives farmers subsidized loans, and the Rural Electrification Administration (REA) makes low-interest loans to rural cooperatives, and guarantees loans for rural telephone and cable television. In 1988, Congress created Farmer Mac (Federal Agricultural Mortgage Corporation) to provide a secondary market for farm mortgages and rural housing loans.[42]

41. See Huber (1989).
42. Farmer Mac is an agency created by the U.S. government and operated by a board of directors, five of whom are appointed by the president of the United States. Like Fannie Mae, Freddie Mac, and Sallie Mae, Farmer Mac is a member of a genre referred to as government-sponsored enterprises, or GSEs. They are hybrid institutions with both private and public aspects. Some have privately owned stock outstanding, but all have some government-appointed board members. Their debts are not explicitly guaranteed by the government, but trade as if there exists some government protection; typically, this ''agency'' debt trades at a spread of less than 50 basis points above Treasury issues of similar duration. The agencies buy, sell, repackage, and guarantee private debts of their constituents, and are exempt from SEC registration requirements.

- **Influencing Credit Allocation Through Consumer Protection Regulation:** Regulation also seeks to ensure that there is no pernicious discrimination by lenders in the allocation of credit. One such regulation is the *Equal Credit Opportunity Act* (ECOA), which is implemented by Regulation B of the Federal Reserve. The ECOA prohibits discrimination on the basis of race, religion, national origin, marital status, age, and gender. The *Home Mortgage Disclosure Act* (HMDA) of 1975 (Regulation C) and the *Community Reinvestment Act* of 1977 (Regulation BB) seek similar ends; see the box below. Regulation C prohibits "redlining," the practice of withholding credit from particular neighborhoods. Regulation BB encourages financial institutions to serve all legitimate credit needs of their communities.

Community Reinvestment Act, 1977

The Community Reinvestment Act (CRA) requires banks to make loans in their own community, even if business judgment calls for deploying the bank's resources elsewhere. Moreover, well before the adoption of the CRA, mechanisms—such as the federal funds market and the Treasury bills market—were already in place to channel funds from deficit to surplus areas. The logic of the CRA is that banks are chartered by the government with privileges and subsidies, and there is a reciprocal obligation to serve the local community where the bank obtains its financial resources (deposits). But as brokers, banks exist to redeploy funds from surplus to deficit users (locales), motivated by spatial differences in interest rates. For example, for decades funds raised via bank deposits on the East Coast found their way to the West Coast because of differences in rates of economic growth and investment opportunities. This welfare-improving transportation of financial resources could violate CRA, depending on how it is applied. The CRA can counterproductively distort credit flows, and this is the source of much of the controversy surrounding this legislation.

Previously, banks with under $250 million in assets (considered "small banks") were mostly tested on whether they were making loans to the entire community. Banks with assets exceeding $250 million (considered "large banks") were tested for lending practices, but were also required to earn 25 percent of their grade in service and 25 percent in community reinvestment. Changes in the CRA rules in 2004 released almost 1,800 banks with assets between $250 million and $1 billion from CRA's data collection requirement as well as community investment and service tests. The new community development test covers four areas of activity: affordable housing, community services, economic development, and revitalization of stabilization activities.

"Redlining" or rejecting credit applicants because of their gender, race, religion, and other attributes is divisive.[43] Moreover, it is wrong to state the cost of a loan or other financial service in terms that intentionally confuse or mislead the client. But it should also be understood that banking is about making the most efficient credit decisions possible on the basis of incomplete information, and gender, race, age, handicap, and neighborhood of origin are correlates of creditworthiness,

43. Recent studies, based on HMDA data, have found evidence of discrimination in lending. See, for example, Cummins (1993a) and Munnell et al. (1992). See also Duca and Rosenthal (1993).

but these attributes are more costly to observe. Thus, the banker need not be a bigot to discriminate. Nor should he or she necessarily be expected to internalize the social costs of using freely available information. But this is the rationale for equal opportunity credit legislation, proponents of which argue that the banker has been privileged to operate with a valuable license, and an obligation attaches to the privilege.

Monetary Control Regulation

The two major forms of bank regulation relating to monetary control objectives are: reserve requirements and the discount window. Moreover, "moral suasion" by the central bank also plays a role. We discuss each in turn.

(1) Legal Reserve Requirements: Cash-asset reserve requirements mandate that the bank retain a certain fraction of its deposits liabilities in a noninterest bearing, liquid form—vault cash or deposits, at the Federal Reserve in the case of Fed member banks, or deposits at member banks in the case of nonmember banks. All depository institutions in the U.S. are subject to reserve requirements on customer deposits.

Reserve requirements also vary depending on the type of the deposit. For instance, reserve requirements against time deposits (CDs) are zero, whereas deposits subject to withdrawal on demand, or net transactions accounts, are subject to a reserve requirement that depends on the level of deposits. No reserves are required for the first \$7 million in deposits. Between \$7 million and \$48.3 million, there is a 3 percent reserve requirement, and above \$48.3 million there is a 10 percent reserve requirement. Reserves are computed as the average held over a 14-day period.

The earliest justification for reserve requirements was as a source of liquidity.[44] This notion was derived from the role that specie reserves played as a source of liquidity to redeem notes. However, required reserves are unavailable to meet deposit withdrawals. The reason is that any deposit withdrawal reduces available reserves, and the deficit must be made up. For example, consider a bank with \$100 in deposits and a 5 percent reserve requirement against these deposits. Imagine that the bank has \$10 in capital, so that the total asset base is \$110. This bank is required to keep \$5 in cash reserves. Imagine that it does so and invests the remaining \$105 in other assets. Now, suppose there is a \$5 deposit withdrawal that the bank meets with its reserves. Since it now has \$95 in deposits, it needs to keep \$4.75 in reserves. This will require taking in new deposits (note, however, that more than \$4.75 in new deposits will be needed since the reserve requirement applies also to the new deposits)[45] or by liquidating other assets.[46]

This is the paradox of fractional reserve requirements. Rather than augmenting liquidity, reserve requirements freeze assets into immobility.[47] The safety of any fractional reserve banking system rests squarely on the availability of a secure and reliable lender of last resort. Fractional reserve requirements cannot help much in this regard.

44. See Edgeworth (1888) and Greenbaum and Thakor (1985).

45. In fact, at least \$5 of new deposits will have to be raised and all of the money invested in cash. If more new deposits are raised, say \$10, a greater amount, \$5.50 in this case, will need to be invested in eligible reserves.

46. For example, the bank may sell some of its marketable securities or loans.

47. Of course, as reserve requirements approach 100 percent, these problems vanish since all of the bank's assets are invested in eligible reserves. This limiting argument is the basis for the once popular 100 percent reserve requirement proposal of Henry Simons (1934, 1935).

More recently, reserve requirements have been rationalized as a tool of monetary policy.[48] In its 1931 report, the Fed Committee on Bank Reserves stated, "The most important function served by reserve requirements is the control of credit." Since increasing reserve requirements means that a smaller fraction of deposits can be loaned out by the bank, the Federal Reserve can, in principle, affect the availability of credit by altering reserve requirements.

As a practical matter, however, reserve requirements have played only a minor role in the Fed's monetary policy. From the early years of the Federal Reserve System through the 1920s, the primary instrument of credit policy was the discount window,[49] and from 1942 until the Treasury-Fed Accord in 1951, reserve requirements remained virtually unchanged because the Fed committed itself to a policy of supporting government bond prices.

The current officially stated rationale for reserve requirements is that they are a tool of monetary policy. This position was first articulated in the 1950s, when the Fed came to view reserve requirements as a mechanism for limiting the growth of the money stock as well as credit. It is now believed, however, that this is a specious argument. Without reserve requirements, banks can be expected to voluntarily hold some cash assets, the amount depending on how the LLR facility is priced and administered, and deposit expansion and contraction would ensue more or less as it would with legal reserve requirements. Moreover, reserve requirements have numerous drawbacks: They foster spurious innovation as depository institutions create deposit substitutes to avoid reserve requirements (since there is invariably a lag before regulators respond by imposing reserve requirements on the new liabilities).[50]

The critical feature of reserve requirements is that they determine the sharing of seigniorage on bank deposits between the central bank and the privately owned banks. The higher the reserve requirement, the greater the share of seigniorage that flows to the Federal Reserve, and ultimately back to the U.S. Treasury. Lower reserve requirements direct these monopoly profits to the privately owned banks. This is why reserve requirements are sometimes referred to as a tax on the banks, but they could be defensibly described as a subsidy, depending on who owns the rightful claim to the deposit seigniorage.

Some have argued that the Federal Reserve continues to support reserve requirements because they produce three bureaucratic benefits.[51] First, reserve requirements permit the remission of substantial sums to the Treasury, thereby fostering the Fed's continued budgetary independence; currently, the Federal Reserve's earnings, after expenses and a small contingency reserve charge, are paid to the Treasury as a special franchise tax.[52] Second, reserve requirements provide the Federal Reserve with a natural constituency since financial institutions subject to reserve requirements can be influenced by the Fed.[53] Finally, in the past, when reserve requirements applied only to Federal Reserve member banks, they enabled the Federal Reserve to expand its

48. For a detailed discussion of the early history of reserve requirements, see Federal Reserve Staff (1938). For a more recent treatment, see Goodfriend and Hargraves (1983).

49. Friedman and Schwartz (1963) argue that the reserve requirement increases of 1936–37 precipitated the economic collapse of 1937–38.

50. See Greenbaum and Higgings (1983), Porter, Simpson, and Mauskopf (1979), Federal Reserve Staff (1979), and Kanatas and Greenbaum (1982).

51. See Greenbaum and Thakor (1985).

52. As of July 2006, required reserves were $45 billion. If the Federal Reserve paid interest at say 1 percent under the discount rate of 6.5 percent at that time, the annual payment to the banks would approximate $2.475 billion.

53. See Kane (1974).

operations because a variety of subsidized services had to be provided to induce members not to leave the system.[54]

The thing to remember about reserve requirements is that their most basic rationale is to address the moral hazard associated with the LLR facility, and the real issue is the sharing of deposit seigniorage between the government and the privately owned banks.

(2) The Discount Window: The discount window is a mechanism by which the Federal Reserve performs its LLR responsibilities. Banks are allowed to borrow through the discount window to meet short-term liquidity needs. Prior to the passage of the Depository Institutions Deregulation and Monetary Control Act of 1980 (DIDMCA), the discount window was available only to member banks. DIDMCA expanded access to the discount window for nonmember banks and S&Ls, mutual savings banks, and credit unions as well. This was "fair" since DIDMCA also extended federal cash-asset reserve requirements to all institutions.

When a depository institution borrows through the discount window, it ordinarily uses government securities as collateral. This borrowing is used to make up reserves lost due to (unanticipated) deposit withdrawals. Thus, the discount window is closely linked to reserve requirements.

Establishment of the discount window was one of the primary reasons for the creation of the Fed. In addition to providing liquidity, the discount window also facilitates the conduct of monetary policy in that contractionary open market operations can drain the liquidity of individual institutions. The box below provides further details on the discount window.

Discount Window Details

- **Brief History:** The Fed has three major instruments of monetary policy: open market operations, changes in reserve requirements, and changes in the terms of borrowing from the discount window. Reserve requirement changes have never been used on a consistent basis as a monetary policy tool. One argument is that they represent a very cumbersome policy instrument, although this is largely a calibration issue. That is, reserve requirements can be changed from 10 percent to 10.0001 percent rather than to 10.5 percent, and then they would not be so cumbersome. Nonetheless, the "cumbersome" argument has often been cited as an impediment to using reserve requirements. And, at least in the early years of the Fed, open market operations were not used much either. Administration of the discount window was the key tool for regulating bank reserves.

  Since one of the stated purposes of the discount window was to encourage bank safety, access to the discount window was considered a privilege rather than an entitlement. At the time that the Fed was created, safety and liquidity were to be promoted by encouraging banks to make short-term, self-liquidating loans backed by real goods ("real bills"). In the early years of the Fed, banks could borrow from the discount window only by discounting eligible commercial paper ("real bills").

(*Continued*)

54. See Gilbert and Peterson (1974).

Banks engaged in risky investments could be denied discount window access by the Fed.

The banking reforms following the Great Depression resulted in the adoption of the principle that banks should be allowed greater access to the discount window. The "real bills doctrine" was discarded, and banks were allowed to borrow at the discount window using any collateral acceptable to the district Reserve Bank. Thus, in the period immediately following the Great Depression, the discount window was used primarily as a means for the Fed to stand ready to act as a lender of last resort and ensure the overall liquidity of the banking system rather than as an instrument of monetary policy or as a way to influence banks to specialize in real bills.[1] In more recent times, the discount window, in conjunction with federal open market operations, has become an important monetary policy tool. For example, when the Federal Reserve wants to stimulate the economy with a monetary expansion, it may lower the borrowing rate at the discount window. This usually has a ripple effect in the economy, lowering a host of other interest rates and facilitating increased borrowing for investment and consumption.

- **The Discount Rate:** The rate at which a depository institution can borrow at the discount window is known as the *discount rate*. This rate is set at each district Federal Reserve Bank by the board of directors and is subject to approval by the Board of Governors of the Federal Reserve System. The idea is for depository institutions to access the discount window only after they have exhausted all other alternatives. The *discount officer* at each district reserve bank has quite a bit of discretion in determining whether to honor a borrowing request.
  The costs of borrowing at the discount window are twofold for a depository institution. One is the *discount rate*, and the other is the cost of the accompanying *increased regulatory surveillance*.
- **Forms of Borrowing From the Discount Window:** There are four forms of discount credit: short-term adjustment credit, extended credit for seasonal purposes, extended credit for other purposes, and emergency credit.
  - **Short-Term Credit:** Short-term credit usually is available for periods less than four weeks and is granted only if all other funding sources have been exhausted. Thus, borrowers cannot be net sellers of federal funds. Moreover, nonbank depository institutions like thrifts must first attempt to obtain credit from their own principal regulator (the FHLBB for S&Ls and the NCUA for credit unions).
  - **Seasonal Credit:** Because some banks, like rural agricultural banks, are subject to predictably large seasonal credit needs and lack ready access to credit markets, the Federal Reserve amended Regulation A in 1973 to provide seasonal credit to banks. This credit is limited to institutions with deposits less than a stipulated maximum, the rationale being that larger institutions have access to credit markets. The credit extension period lasts from four weeks to nine months.
  - **Other Extended Credit:** This credit is available for institutions that are suffering from protracted liquidity problems and is intended to give an institution "breathing room" to recover from loan losses or other liquidity problems. DIDMCA also made extended credit available to thrifts experiencing disintermediation due to interest-rate shifts. Extended credit use has risen sharply in the last two decades, as banks and other depository institutions have experienced more severe liquidity problems. In 1984, for example, extended credit was used in the rescue of Continental Illinois National Bank and Trust Company of Chicago. However,

FDICIA made the Federal Reserve liable for certain losses to the Bank Insurance Fund of the FDIC when a bank forestalls failure through discount window borrowings but later fails anyway.[2]

- **Emergency Credit for Others:** Emergency credit is available to individuals and businesses. Such credit is extended only under very rare circumstances by the district Federal Reserve Bank, and only after consultation with the Board of Governors. Such credit is ostensibly made available only if the borrower is unable to secure credit elsewhere, and failure to obtain credit could have a harmful effect on the economy.

1. The Fed may still utilize the discount window to modify the behavior of potential users. For instance, in 1966 the Fed discouraged member banks from making certain types of business loans, and those who cooperated were assured easier access to the discount window. It is not clear how quantitatively important the discount rate has been. Changes in borrowed reserves seem to be only marginally influenced by changes in the discount rate. Discount rate changes may, however, have a bigger role to play in changing expectations about the future, that is, as a signaling device.

2. Mandated by FDICIA in 1991, this rule was approved by the Federal Reserve in December 1993. See Cummins (1993b).

(3) Moral Suasion: Central banks around the world also exercise control over banks they regulate by using "moral suasion" or "jawboning." This is simply exerting pressure on banks by persuading, cajoling, or coercing them to act in a particular way. This mode of policy implementation is less feasible in the United States with its thousands of banks than in Europe or Japan where banking tends to be more concentrated.

Conclusion

Banks have been regulated for over two centuries in the United States. Although regulation has been shaped largely by historical events, as opposed to being the outcome of a well-thought-out regulatory agenda, there have been some important goals that have guided banking regulation. In this chapter, we have explained these goals and described the major regulations to which banks are subject. In the next chapter, we discuss important milestones in banking legislation in the 1990s, and the early 21st century.

Review Questions

1. What are the main objectives of bank regulation? Discuss each.
2. How inherent is the need for bank regulation? Relate your answer to the *raison d'être* for banks.
3. Which are the main agencies of bank regulation, and what is the function of each?
4. Why do we have reserve requirements? What are their drawbacks?
5. What is the purpose of the discount window?
6. Why do we have capital requirements? What are the components of a good capital standard?

7. Critique the Basel I Accord on internationally harmonized capital standards.
8. Discuss the key elements of safety regulation in banking. What specific role does each play in ensuring bank safety? To what extent are these regulations complements or substitutes in this regard?
9. Discuss the economics of branching and bank holding company legislation in the United States.
10. What impact has vagueness in bank holding company legislation had on the behavior of U.S. banks?
11. Discuss the division of a bank's capital into Tier-1 and Tier-2 capital. Contrast this with the usual definition of capital in a nonfinancial firm. Why do you think banks have this more elaborate definition of capital and a division of capital into Tier-1 and Tier-2 components?

References

Acharya, Viral, and Tanju Yorulmazer, "Too Many to Fail: An Analysis of Time Inconsistency in Bank Closure Policies," *Journal of Financial Intermediation* 16, January 2007, 1–31.

Bacon, Kenneth H., "FDIC Bank Fund Soars to $6.8 Billion, But Bankers Claim That's Understated," *The Wall Street Journal*, August 11, 1993.

Benston, George J., and George G. Kaufman, *Risk and Solvency Regulation of Depositor Institutions: Past Policies and Current Options*, Monograph Series in Finance and Economics, New York University, 1988.

Besanko, David, and Anjan V. Thakor, "Banking Deregulation: Allocational Consequences of Relaxing Entry Barriers," *Journal of Banking and Finance* 16–5, September 1992, 909–932.

Besanko, David, and George Kanatas, "Bank Capital Regulations: Time Consistency, Hedging, and Incentive Compatibility," unpublished manuscript, 1991.

Bhattacharya, Sudipto, Arnoud Boot and Anjan V. Thakor, "The Economics of Bank Regulation," *Journal of Money, Credit and Banking* 30–4, November 1998, 745–770.

Boot, Arnoud W.A., and Stuart I. Greenbaum, "American Banking Legislation, Recent," in *The New Palgrave Dictionary of Money & Finance*, editors, Peter Newman, Murray Milgate, and John Eatwell, The Macmillan Press Ltd., New York, 1992, 35–37.

Boot, Arnoud W.A., Stuart I. Greenbaum, and Anjan V. Thakor, "Reputation and Discretion in Financial Contracting," *American Economic Review* 83–5, December 1993, 1165–1183.

Buser, Stephen, Andrew Chen, and Edward Kane, "Federal Deposit Insurance Regulatory Policy and Optimal Bank Capital," *Journal of Finance* 36, 1981, 51–60.

Coheo, Steve, "Fed Adopts Discount Window Rules Meant to Limit Sick Banks' Credit," *American Banker*, December 16, 1993b.

Cummins, Claudia, "Nader Study Accuses 49 Lenders of Redlining," *American Banker*, August 13, 1993a.

Duca, John V., and Stuart S. Rosenthal, "Borrowing Constrains, Household Debt and Racial Discrimination in Loan Markets," *Journal of Financial Intermediation* 3–1, October 1993, 77–103.

The Economist, May 8–14, 1993.

Edgeworth, Francis Y., "The Mathematical Theory of Banking," *Journal of the Royal Statistical Society* 51, 1888, 113–127.

Federal Reserve Staff, "Redefining Monetary Aggregates," *Federal Reserve Bulletin* 65, 1979, 13–42.

Federal Reserve Staff, "The History of Reserve Requirements for Banks in the United States," *Federal Reserve Bulletin* 24, 1938, 953–972.

Fisher, Gerald C., *The Modern Bank Holding Company: Development Regulation, and Performance*, Temple University, Philadelphia, PA, 1986.

Flannery, Mark J., "Removing Deposit Rate Ceilings: How Will Bank Profits Fare?" *Economic Review*, Federal Reserve Bank of Philadelphia, March/April 1983, 13–24.

Friedman, Milton, and Anna J. Schwartz, *A Monetary History of the United States: 1867–1960*, National Bureau of Economic Research, 1963.

Gasbarro, Dominic, I Gde Made Sadguna, and Kenton J. Zumwalt, "The Changing Relationship Between CAMEL Ratings and Bank Soundness During the Indonesian Banking Crisis," *Review of Quantitative Finance and Accounting* 19, 2002, 247–260.

Gennotte, Gerard, and David Pyle, "Capital Controls and Bank Risk," *Journal of Banking and Finance* 15, 1991.

Gilbert, Gary G., and Manfred O. Peterson, "Uniform Reserve Requirements on Demand Deposits: Some Policy Issues," *Journal of Bank Research* 5, Spring 1974, 38–44.

Golembe, Carter, "The Deposit Insurance Legislation of 1933: An Examination of Its Antecedents and Purposes," *Political Science Quarterly* 75, 1960, 181–200.

Goodfriend, Marvin, and Monica Hargreaves, "A Historical Assessment of the Rationales and Functions of Reserve Requirements," *Economic Review* 2, Federal Reserve Bank of Richmond, March/April 1983.

Greenbaum, Stuart I., and Anjan V. Thakor, *Legal Reserve Requirements in Banking: A Review and Assessment*, monograph, Bank Administration Institute, 1985.

Greenbaum, Stuart I., and Bryon Higgins, "Financial Innovation," in *Financial Services: The Changing Institutions and Government Policy*, G.J. Benston, editor, The American Assembly, Columbia University, Englewood Cliffs, NJ: Prentice Hall, Inc., 1983, 213–214.

Greider, William, *Secrets of the Temple*: *How the Federal Reserve Runs the Country*, Simon and Schuster, New York, 1987.

Huber, Stephen K., *Bank Officer's Handbook of Government Regulation*, Boston: Warren, Gorham & Lamont, Inc., 1989.

Kahane, Yehuda, "Capital Adequacy and the Regulation of Financial Intermediaries," *Journal of Banking and Finance* 1, June 1977, 207–218.

Kanatas, George, and Stuart I. Greenbaum, "Bank Reserve Requirements and Monetary Aggregates," *Journal of Banking and Finance* 6, December 1982, 507–520.

Kane, Edward, "All for the Best: The Federal Reserve Board's 60th Annual Report," *American Economic Review* 64, December 1974, 842–850.

Koehn, M., and Anthony M. Santomero, "Regulation of Bank Capital and Portfolio Risk," *Journal of Finance* 35, 1980, 1235–1250.

Lash, Nicholas A., *Banking Laws and Regulations: An Economic Prospective*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1987.

Merton, Robert C., "Influence of Mathematical Models in Finance on Practice: Past, Present and Future," WP #94–045, Division of Research, Harvard Business School, January 1994.

Mester, Loretta J., "Banking and Commerce: A Dangerous Liaison?" *Federal Reserve Bank of Philadelphia Business Review*, May/June 1992, 17–29.

Munnel, Alicia H., Lynn E. Browne, James McEneaney, and Geofferey M.B. Tootell, "Mortgage Lending in Boston: Interpreting HDMA Data," working paper 92–7, Federal Reserve Bank of Boston, October 1992.

Porter, Richard D., Thomas D. Simpson, and Eileen Mauskopf, "Financial Innovation and Monetary Aggregates," *Brookings Papers on Economic Activity*, 1979, 213–229.

Repullo, Rafael, "Capital Requirements, Market Power, and Risk-Taking in Banking," *Journal of Financial Intermediation* 13, 2004, 156–182.

Reuters, "All of World's Big Banks Said to Meet Risk-Based Rules," *American Banker*, April 26, 1993.

Robins, Philip K., "The Effect of State Usury Ceilings on Single Family Homebuilding," *Journal of Finance*, March 1974, 227–231.

Rose, Harold, "Bank Risk and the Level Playing-Field," Working Paper 166, London Business School Institute of Finance and Accounting, 1993.

Simons, Henry C., "Rules Versus Authorities in Monetary Policy," *Journal of Political Economy*, February 1935, 1–30.

——, *A Positive Program for Laissez Faire*, Public Policy Pamphlet No. 15, University of Chicago, 1934.

Thakor, Anjan V., and Jess Beltz, "A Barter Theory of Bank Regulation and Credit Allocation," forthcoming, *Journal of Money, Credit and Banking*, 1994.

Von Thadden, Ernst-Ludwig, "Bank Capital Adequacy Regulation Under the New Basel Accord," *Journal of Financial Intermediation* 13, 2004.

CHAPTER ♦ 12

Milestones in Banking Legislation and Regulatory Reform

"In all that the people can individually do well for themselves, government ought not to interfere."

Abraham Lincoln

Glossary of Terms

BHCA: Bank Holding Company Act.

Golden Parachute: A severance payment made to a manager upon termination of employment.

Sarbanes – Oxley Act: Legislation that mandated more restrictive corporate governance guidelines for publicly traded companies.

TBTF: Too Big to Fail.

GAO: General Accounting Office.

Narrow Bank: A bank that is restricted in its assets. The original narrow bank proposal called for all of the deposits to be invested in cash and government securities.

LLR: Lender of Last Resort.

Universal Bank: A financial intermediary that performs services usually associated with commercial banks, investment banks, and insurance companies.

Introduction

In this chapter, we discuss milestones in banking legislation and review bank regulatory reform proposal. This chapter is a continuation of Chapter 11, which looked at the objectives of regulation. Here we examine what has happened and what lies ahead. First, we will describe legislative milestones in banking. Then we will review major problems of bank regulation, after which we examine the causes of and possible cures for these problems. Next, we turn to the Federal Deposit Insurance Corporation Improvement Act (FDICIA) of 1991, and the Basle II Capital Accord adopted in 2004. Regulation clearly seems to be shifting toward greater reliance on effective capital requirements for banks, but the traditional view of how monetary policy works assumes that cash-asset reserve requirements are binding. This issue is examined just before the conclusion of the chapter.

Milestones of Banking Legislation

Banking legislation has shaped the relationship between government and privately owned banking institutions from the earliest history of the United States. The first banks were chartered by the states, but the federal government reserved to itself the control of interstate commerce and the production of coin and currency.

With growing governmental responsibility for stabilizing economic activity came increasing involvement with the banks. Failures and financial panics linked to banks preceded recessions, and many believed that banks were instrumental in producing financial panics and business cycles. Six major eras of U.S. banking regulation are summarized in Figure 12.1.

Early Bank Regulation

Eighteenth and Nineteenth Century Banking: The creation of the *Bank of North America* in 1781 was driven by the fledgling government's need for a fiscal agent. Soon after the colonies won their independence, the Continental Congress gave a perpetual charter to the Bank of North America. Later, other banks emerged. However, criticism of lending policies and the ability to issue paper currency led to a repeal of the charter given to the Bank of North America.

The First and Second Banks of the United States: With the active support of Alexander Hamilton, then Secretary of the Treasury, the First Bank of the United States was chartered in 1791 for a 20-year period. The bank was an embryonic central bank in that it issued notes, accepted deposits, transferred government funds through its eight branches, made public disbursements, and granted credit to the government as well as the private sector. However, the First Bank did not serve as a depository of bank funds, or as a clearinghouse, or as a creator of bank reserves. Nor did it act as an LLR. The bank was severely criticized for its "anti-South" bias, its inattention to agrarian interests, and its growing English ownership. With Congress not renewing its charter, the bank expired in 1811.

In 1816, Congress chartered the Second Bank of the United States. This bank, initially a fiscal agent for the government, evolved into an embryonic central bank. It would, for example, redeem the bank notes of suspect institutions. The bank was

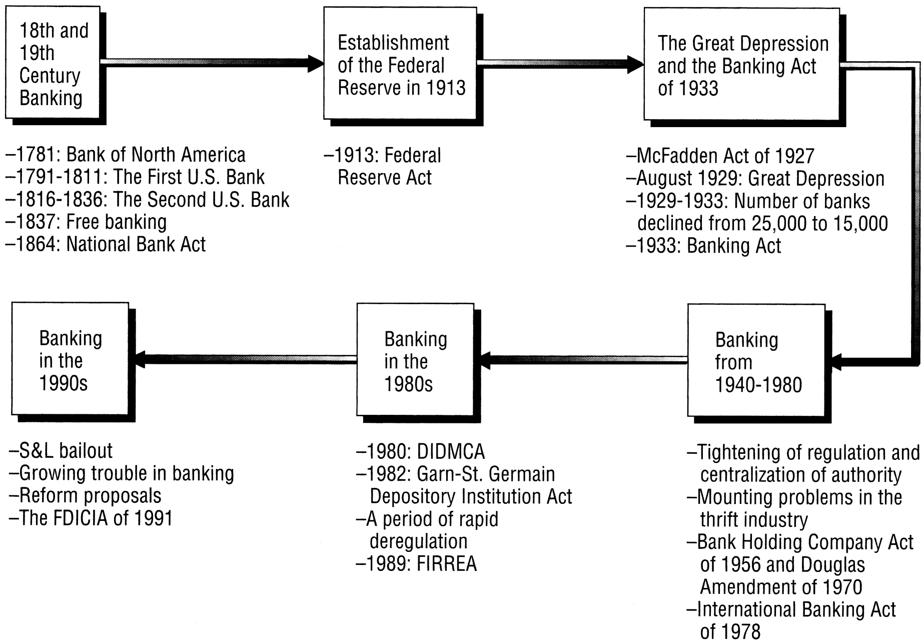


FIGURE 12.1 The Major Eras of U.S. Bank Regulation

seen, however, as a disciplinary agent representing eastern (lenders) interests at the expense of agrarian (borrowers) interests. When Andrew Jackson, representing the agrarian and frontier interests, was elected president, efforts to recharter the Second Bank of the United States were stifled. Its federal charter expired in 1836.

The period from 1837–64 is commonly referred to as the era of *free banking*. It was a period of minimal federal government involvement in banking. The states had virtual free rein. Colorful stories of "wildcat banking" circulated along with bank notes of heterogeneous value. These banks would open in remote locales in order to frustrate note redemption efforts.

National Bank Act of 1864: The National Bank Act of 1864 marked the return of the federal government to banking. With the 5 percent tax on state bank notes and licensing of national banks, the era of free banking was brought to a close. The National Bank Act established the *Office of the Comptroller of the Currency* (OCC) to charter and supervise national banks and to regulate the national currency. With the tax on state bank notes, the largest and most reputable banks obtained national bank charters.

The National Bank Act probably had more to do with financing the Civil War than with reforming banks; national banks were required to hold government securities to satisfy liquidity requirements. Moreover, the populist distrust of banks, which sought to avoid undue concentration of power, led to a fragmented banking industry structure.

The period following the Civil War was characterized by periodic financial disruptions as banks' liquidity would be tested by skittish note and deposit holders. Inevitably, some banks would be found wanting and contagious panics would occasionally ensue. Systemic risk arose from provisions that allowed banks to hold their reserves in the form of deposits at other banks. This pyramiding of reserves and attendant panics eventually led to the creation of the Federal Reserve System in 1913. In addition, the pre-Federal Reserve monetary system was inefficient with the notes of thousands of individual banks circulating as imperfect substitutes for one another. This was an era of not only private deposits, but also privately produced currency. There were as many media of exchange as there were banks.

The Federal Reserve Act: Following a particularly disruptive financial panic in 1907, the Congress created the National Monetary Commission to recommend reform of the banking system. Their work led to the Federal Reserve Act in 1913 that established the Federal Reserve System.

The United States was the last major western country to establish a central bank. Unique in its decentralized design, the Federal Reserve reflected the historical ambivalence about creating a powerful quasi-government banking institution. America's deep-rooted populism recoiled at the notion of a centralized hegemony over banking. The genius of the system is that it has been able to function credibly despite its convoluted design. Nominally privately owned, it is a governmental institution. Nominally decentralized with 12 separate corporate entities, virtually all important decisions are made by the presidentially appointed Board of Governors in Washington, D.C.

Originally, the Federal Reserve had note-issuing authority, LLR powers, and performed clearing services. But with time, the Federal Reserve took on increasing responsibilities for monetary policy and bank regulation. Perhaps most important among the latter are its responsibilities for oversight of bank holding company activities.

Legislation During 1920–1980

The McFadden Act of 1927: The McFadden Act addressed the question of geographic expansion of national banks. Each of the states retained the power to determine the basis on which state-chartered banks could expand their facilities or branches. Thus, states like Illinois limited banks to having only one office (hence the term unit banking). Others, like California, placed no limits on the branching powers of their banks; California banks could establish offices anywhere in the state. Still others like New York permitted limited-area branching. Similarly, some states permitted multibank holding companies whereas others explicitly forbade bank holding companies.

A question arose about the powers of national banks. Prior to McFadden, some state banks had more expansive branching powers than competing national banks. The McFadden Act gave national banks exactly the same powers as state banks in the states where the national banks are domiciled. Thus, national banks domiciled in California would have the same branching powers as California's state-chartered banks, and national banks located in Illinois would be restricted in the same way as Illinois' state-chartered banks.

This principle of devolution preserved the dual banking system and the fragmentation of banking markets.[1] While it put state and national banks on an equal footing, it also prevented national banks from expanding nationwide, thereby limiting banks' ability to diversify their funding and credit risks and to exploit economies of scale. This diversification problem would come back to plague the industry in the 1980s.

1. Later legislation overrode state restrictions on branching for thrifts. For a detailed account of early American banking history, see Hammond (1957).

The distress and failure of Texas' major banks in the 1990s was in good part due to their undiversified exposure in energy-related industries. And Texas was not alone. The energy industry funk brought down major banks in Oklahoma, Louisiana, and Colorado, too. The subsequent difficulty of New England banks was similarly linked to a regional recession. And still more recently, cutbacks in defense spending and other local problems have stressed California's banks.

The McFadden Act and its litigation of interstate bank expansion was eventually mooted by the Gramm-Leach-Bliley Act, which permitted well-capitalized banks to expand across state lines. Even with the rapid decline in the number of banks and complementary increase in market share of the top-10 banks there is still a paucity of banks in the U.S. with facilities that are distributed nationwide.

Glass-Steagall Act of 1933: From 1919 to 1929, 6,000 banks were suspended or liquidated, and another 4,000 merged with other banks. From 1929 to 1933, another 10,000 banks failed, as the number of banks declined from 25,000 to 15,000. In the depths of the nation's worst economic recession (GNP dropped by 50 percent, the money supply fell by 33 percent and the unemployment rate reached 25 percent in 1932), failing banks were a focal point of discontent. There was no more potent force transforming ordinary folk into revolutionaries than the loss of one's liquid assets in some ostensibly mismanaged bank. (Recall there were no readily available risk-free assets other than currency. Mutual funds came much later and government securities were available only in large denominations.)

This was the ambiance in which newly elected President Franklin D. Roosevelt set out to reform banking. The crisis was memorialized with the Bank Holiday of March 1933 that closed all banks. Congress then shaped legislation that ultimately reconfigured banking more fundamentally than any previous legislation in U.S. history.

The 1933 legislation introduced federal deposit insurance despite President Roosevelt's misgivings. This reform addressed the public's need for a risk-free asset and stemmed the flight from bank deposits to currency. The legislation also capped deposit interest rates, providing banks with a new subsidy. Together with the follow-on Banking Act of 1935, Glass-Steagall took banks out of the securities business and imposed more intrusive supervision than ever before.[2] The most significant provisions of the 1933 legislation are summarized in Figure 12.2.

This legislation was lauded as one of the most successful governmental intrusions into the private sector, ever. President Roosevelt's misgivings about deposit insurance took a full 50 years to be realized. Until the inflation of the 1970s and 1980s, the premium charged for deposit insurance was less than 1/10 of 1 percent of the deposit base per year. The deposit insurance fund grew steadily, and bank failures were inconsequential. Glass-Steagall, and more particularly the deposit insurance it established, was one of the most admired monuments of the New Deal. As pointed out in Chapters 10 and 11, the remarkable point about deposit insurance is not that it eventually came unraveled, but rather that it lasted as long as it did. President Roosevelt clearly foresaw the moral hazards in the deposit insurance system. What he could not be expected to understand was that these internal contradictions could be held in check for four decades.

Two years after Glass-Steagall, the Banking Act of 1935 became law. It renamed the Federal Reserve Board as the *Board of Governors of the Federal Reserve System* and extended its powers to regulate the discount rates of the district Federal Reserve

2. See Benston (1990).

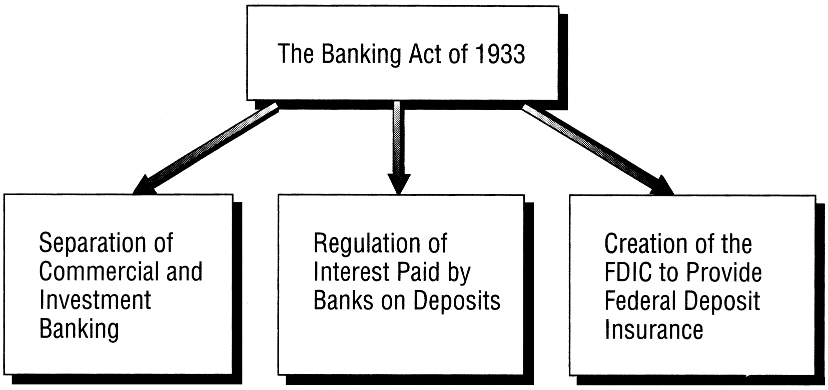


FIGURE 12.2 Significant Provisions of the Banking Act of 1933

banks and cash-asset reserve requirements, and impose margin requirements on securities lending. The committee coordinating open-market operations was renamed the *Federal Open Market Committee* (FOMC).

Banking During 1940–1980: The U.S. banking system came out of World War II with immense holdings of U.S. government securities and cash-asset reserves. Banks were well capitalized and credit risk was a minor problem. Indeed banks' loan portfolios had grown very little during the greatest wartime mobilization in U.S. history. Much of the build-up was directly financed by government with the banks serving the secondary role of accumulating government debt. (Recall that the banking industry was still emerging from the trauma of the Great Depression.)

All of this changed when the widely predicted postwar economic funk failed to materialize. Pent-up demand of returning veterans unleashed a sustained prosperity, and bankers sought ways to participate. This required re-examination of their bomb-shelter mentality and the development of methods to prudently process greater risk, principally credit and liquidity risk.

Thus, the postwar period saw banks that had emerged from World War II with over 74 percent of their assets in government securities replace government securities with business loans. Balance sheets grew, capital ratios fell, and so did cash-asset reserve ratios. It was in this climate that banks expanded their branch systems and began holding company powers more aggressively.

Bank Holding Company Act (BHCA) of 1956 and the Douglas Amendments of 1970: Although group (holding company) banking grew little from 1933 to 1948, activity picked up considerably from 1948 through 1956.[3] Concern about the use of the BHC to expand geographically and functionally prompted the BHC legislation of 1956 and the Douglas Amendments of 1970. Prior to this legislation, the federal government had little power to regulate or supervise BHCs, the corporate parents of the banks. This was viewed as a loophole that needed to be addressed. The BHCA defined a

3. See Fisher (1986). Legislation designed to subject BHCs to stricter regulation had been introduced at every session of Congress between 1933 and 1955, so there was concern among bankers about stricter margin restrictions. This fear of pending legislation prompted a rapid development in multiple-unit banking.

BHC as any entity that owns or controls 25 percent or more of the voting shares and controls the board of directors of two or more affiliated banks. The 1956 law required BHCs to: (i) divest ownership of businesses other than banking or furnishing services to affiliated banks, (ii) register with the board of Governors of the Federal Reserve, and (iii) seek approval of the board for any bank acquisitions.

Although the Federal Reserve was charged with primary responsibility for regulating BHCs, the focus was on bank acquisitions of holding companies and multibank holding companies. Thus, the 1956 legislation largely ignored questions raised by nonbank acquisitions of one-bank holding companies. This was because nonbank acquisitions were not yet an issue in 1956. But the relentless testing of the limits of banking, symptomatic of the segue away from banking's depression mentality, brought this latter issue to the fore. The 1970 Douglas Amendments required all BHC acquisitions to have explicit Federal Reserve approval. The Federal Reserve developed a laundry list of approvable and prohibited activities, but these lists were merely presumptive, and each individual acquisition required explicit approval. The Douglas Amendment's charge to the Federal Reserve was ambiguous—"[nonbank acquisitions] should be so closely related to banking to be a proper incident thereto"—and the Federal Reserve consequently has virtually boundless discretion in deciding on BHC acquisition applications.

The importance of the BHC regulation, both extant and prospective, is clear in the current debate on reform. First, existing legislation clearly lodges almost boundless power in the Federal Reserve. Second, virtually all proposals to expand banking powers rely on the holding company and its questionable "fire walls" to protect the bank and its insured deposits. Almost certainly, the role of the BHC will expand as banking legislation is liberalized, and the Federal Reserve will be the regulatory focal point.

International Banking Act of 1978: The International Banking Act of 1978 was designed to provide a more "level playing field" between U.S. banks and their foreign-bank competitors operating in the U.S. market. Foreign-bank branches were compelled to select one state as domicile for McFadden purposes. They also were required to satisfy capital and liquidity requirements comparable to those of their U.S. competitors.

Incidental to this complex exercise in defining equivalence came the first explicit continuous capital requirement for banks.[4] This was another testimony to the success of the 1930s legislation. The 5.5 percent capital requirement of the 1978 legislation was almost an afterthought to defining equivalence between U.S. and foreign banks competing in U.S. markets.

Problems of the Thrift Industry: The 1970s saw a significant increase in interest rate levels and volatility owing to high and volatile inflation rates. In addition, information technology improved dramatically. These two developments profoundly affected banks and thrifts. Interest-rate surprises led to crippling losses for financial institutions with mismatched balance sheets. The thrifts that were legally locked into long-term, fixed-rate mortgages suffered worse than banks that had gradually substituted floating-rate loans for fixed-rate term loans. Advances in information technology

4. Previously banks had to satisfy a minimal capital requirement at their moment of birth, but absent insolvency little was said about capital thereafter.

weakened barriers to entry and invited competition from a wide variety of nonbank providers of financial services, such as mutual funds, finance companies, and the capital markets. By 1980, the thrift industry was on the brink of insolvency, and banking failures were increasing in size as well as frequency.

Legislation of the 1980s and the 1990s

Depository Institutions Deregulation and Monetary Control Act (DIDMCA) of 1980: DIDMCA addressed two major issues: the disintermediation of deposits that was exacerbated by deposit interest-rate ceilings, and the attrition of Federal Reserve membership as more banks sought to avoid the cost of maintaining cash-asset reserve requirements.

Deposit interest-rate ceilings introduced by the Glass-Steagall legislation had always been a mixed blessing for the banks. When market interest rates for deposit substitutes—government securities, money market mutual funds—were only moderately higher than the ceilings, the banks benefited owing to depositor inertia or convenience. But as the disparity between market and ceiling rates widened, depositors became restless and funds flowed out of the banks. The same deposit interest ceilings that were a major support of banking when interest rates were tranquil became a headache as interest rates became more volatile.

This problem might have been addressed by indexing the deposit interest-rate ceilings to market interest rates, but this was never done. The regulators seemed to prefer unlimited discretion, but their efforts to make timely adjustments in the ceilings could not keep up with the fast moving capital markets.

DIDMCA addressed this problem by providing for the gradual elimination of all deposit interest-rate ceilings, except those on demand deposits.[5] Banks became free to compete in deposit markets as they saw fit. Even the demand deposit interest-rate restriction was circumvented with consumer NOW accounts. What was not too clearly understood was how important the earlier deposit subsidies had been in discouraging high-risk strategies of banks. DIDMCA also raised the ceiling on federally insured deposits from $40,000 to $100,000 per account. This reduced the incentive of depositors to monitor their banks, further encouraging risk-taking by banks.

The second major initiative of DIDMCA was to subject all insured banks to Federal Reserve cash-asset reserve requirements. This addressed the Federal Reserve's problem of membership attrition. The opportunity cost of satisfying the Federal Reserve's cash-asset reserve requirements increased with the level of market interest rates and declining Federal Reserve membership was yet another piece of the syndrome of the 1970s. Thus, DIDMCA eliminated an opportunity for regulatory arbitrage. Interestingly, since 1980 the Federal Reserve has lowered demand deposit reserve requirements from a maximum of over 16 percent to 10 percent. Thus, regulatory hegemony has been accompanied by a transference of the taxpayer's seigniorage to the banks. A summary of the major provisions of DIDMCA appears in the box below.

5. One catalyst for DIDMCA was a Supreme Court deadline for addressing the alleged illegality of NOWs, ATMs, and share draft accounts.

Major Provisions of DIDMCA

- All depository institutions were permitted to issue interest-bearing checking accounts and required to hold cash-asset reserves as prescribed by the Federal Reserve.
- S&Ls were allowed to have up to 20 percent of their assets in a combination of consumer loans, commercial paper, and corporate debt instruments.
- Federal S&Ls were allowed to offer credit-card services and engage in trust activities.
- A statutory capital requirement for S&Ls of 5 percent of deposits was replaced with a range of 3 to 6 percent to be set by the Federal Home Loan Bank Board.
- Deposit interest-rate ceilings were phased out over a 6-year period. Interest-rate deregulation was to be administered by the Depository Institutions Deregulation Committee (DIDC) with the Secretary of Treasury as chair and the heads of the Federal Reserve, the FDIC, the Federal Home Loan Bank Board, and the National Credit Union Administration as voting members.
- The deposit insurance limit was raised to $100,000 per account.
- Statewide branching was permitted for federal S&Ls.
- Earlier geographical limits on S&L lending—loans could only be made within a 50-mile radius of an office—were eliminated.
- Authority of federal S&Ls to make acquisition, development, and construction (ADC) loans was expanded.

Garn-St. Germain Depository Institutions Act of 1982: The Garn-St. Germain Act was directed at thrifts exclusively and sought to enhance their earnings potential by expanding their powers. The initiative was a response to the huge losses suffered by the industry due to the 1980–81 spike in interest rates.

The prime rate soared to over 21 percent. The thrift industry was forced to fund its vast portfolio of loans and fixed-rate mortgages with very high cost liabilities. The losses sustained over an 18-month period eroded a significant portion of the industry's capital.[6] Much of the loss was attributable to thrifts having been legally confined to fixed-rate mortgages. In order to earn their way back, it was argued that thrifts needed more liberalized asset empowerments, including the authority to make adjustable-rate mortgages.

Garn-St. Germain provided the expanded asset powers the industry sought. This permitted vastly increased credit risk, and those who had been most devastated by losses were the most eager to pursue high-risk strategies. A summary of the major components of Garn-St. Germain is given in the box below.

6. Estimates range between $150 billion [Balderstone (1985)] and $165 billion [Kane (1990a)]. See also Kane and Yu (1993). For perspective, note that thrifts were a trillion-dollar industry with something less than 5 percent in capital.

Major Provisions of Garn-St. Germain

- Asset powers of federal S&Ls were expanded by permitting:
  Up to 40 percent of assets in commercial mortgage loans.
  Up to 30 percent of assets in consumer loans.
  Up to 10 percent of assets in commercial loans.
  Up to 10 percent of assets in commercial leases.
- Elimination of the previous statutory limit on the loan-to-value ratio, allowing S&Ls to lend more relative to the appraised value of a project.
- Authorization of the FDIC and the FSLIC to issue "net worth certificates" that could increase an institution's capital for regulatory purposes without any real infusion of capital.

Financial Institutions and Regulatory Reform Enforcement Act (FIRREA) of 1989: FIRREA was the sequel to Garn-St. Germain. It created the machinery and procedures to dispose of insolvent and near-insolvent thrifts. The regulatory agency for thrifts, the FHLBB, was disenfranchised. A new thrift regulator was created, the Office of Thrift Supervision (OTS), within the Treasury Department. The thrift insurer, the FSLIC, was also reorganized and placed within the FDIC as the Savings Association Insurance Fund (SAIF). The legislation also created the Resolution Trust Corporation (RTC) to dispose of failed thrifts and their assets.

FIRREA laid the groundwork for more aggressive resolutions of impaired thrifts, and subsequently more than 1,000 thrifts were restructured via government intervention. FIRREA sought to correct for the passivity and forbearance of earlier regulatory policies. The major elements of FIRREA are outlined in Figure 12.3.

FIRREA had two important provisions insofar as capital requirements are concerned. First, FIRREA stipulated three types of capital requirements: tangible capital, core capital and risk-based capital. Tangible capital is common equity and perpetual preferred stock; the OTS required thrifts to keep tangible capital equal to at least 1.5 percent of total assets. Core Capital was defined as tangible capital plus nonperpetual preferred stock and qualifying subordinated debt; the OTS required thrifts to keep core capital equal to at least 3 percent of total assets. Intangible assets like goodwill could no longer count as part of core capital by 1994, with a phase-out schedule stipulating the rate at which intangible assets had to be phased out from consideration as regulatory capital. With the passage of FIRREA, the term "supervisory goodwill" was used to denote goodwill created in FSLIC-assisted acquisitions of insolvent thrifts in which there was a specific agreement by regulators to permit the goodwill to count as regulatory capital. It is this goodwill that was subject to a 5-year phaseout. All other goodwill was immediately disqualified as regulatory capital. The risk-based capital ratio had to exceed 8 percent and was to be computed in the same way as the Tier-1 capital ratio under the Basel Accord.

Banking in the 1990s: FIRREA mandated a study by the Treasury that would propose reform of the deposit insurance system. This resulted in a February 1991 proposal by

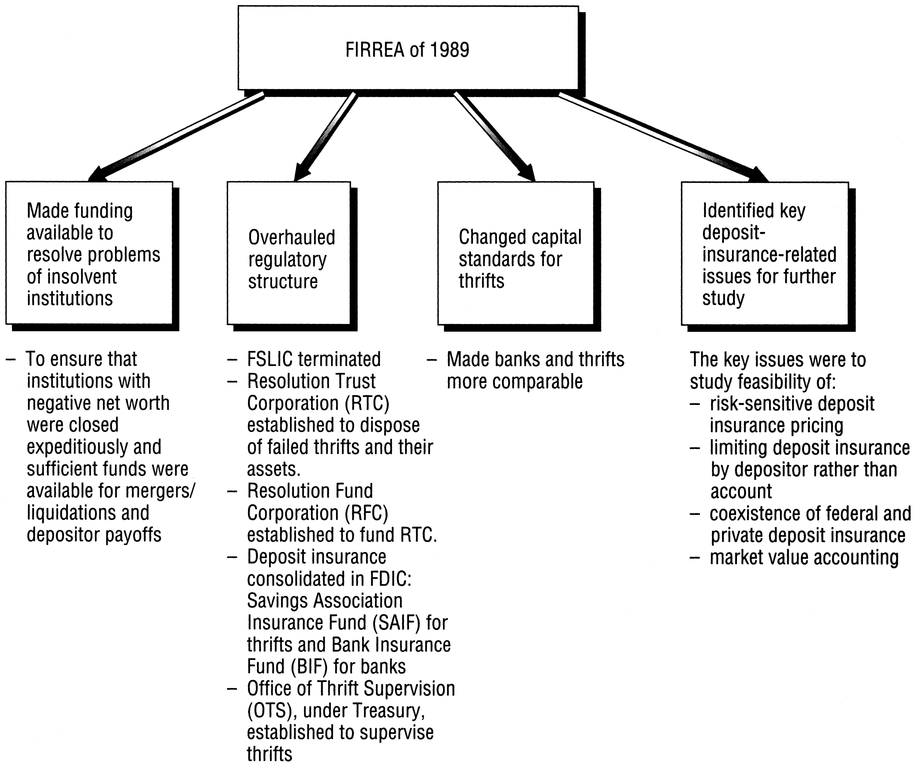


FIGURE 12.3 Major Elements of FIRREA

the Treasury. The Treasury's bold reform proposals were rejected in favor of the FDICIA that we discuss later in this chapter.

Problems of Bank Regulation

The Problems

In Table 12.1 we summarize prominent problems among deposit-taking financial institutions, along with possible causes and commonly suggested remedies. Our perspective here is that of a taxpayer desiring a globally competitive banking system in which taxpayer exposure is minimized.

Causes and Possible Cures

In this subsection, we briefly discuss each of the major issues and possible remedies listed in Table 12.1. Our discussion is organized around: deposit insurance, regulatory uncertainty, market value accounting, and expanded banking powers.

TABLE 12.1 Symptoms, Causes and Commonly Suggested Cures for the Problems of Depository Institutions in the 1980s

Symptoms	Possible Causes	Commonly Suggested Cures
• Excessive risk-taking. • Management fraud.	–Distorted incentives arising from the pricing of deposit insurance. –Ineffective regulatory monitoring. –Low charter values of depository institutions.	–Improving capital standards. –Risk-sensitive deposit insurance. –Restricted entry into banking. –"Narrow" banks. –Improving monitoring procedures. –Providing greater resources for regulatory surveillance.
• Excessive delays in dosing failed institutions—forbearance.	–Regulatory accounting principles (RAP). –Self-interested bank regulators.	–Market value accounting. –Improving incentives of regulators.
• Unpredictable effects of monetary policy.	–Reserve requirements. –Financial innovation.	–Modify reserve requirements. –Reduce regulatory taxes that encourage financial innovation.
• High cost of equity capital for banks.	–Regulatory uncertainty.	–Make regulation more predictable and eliminate perceived regulatory capriciousness.
• Declining competitiveness of U.S. banks.	–Improved information processing in economy and reduced value of banking services. –Loss of market share to foreign banks.	–Less onerous regulation. –Expanded powers for banks to permit entry into investment banking and insurance. –Dismantling of branching restrictions. –International harmonization of capital standards.

• **Deposit Insurance and the Bank's Incentives:** Our discussion in Chapter 10 highlighted problems associated with deposit insurance. It is commonly believed that these problems arise from: (i) the pricing of deposit insurance, (ii) the incentives of regulators, and (iii) the incentives of bank executives. We discuss each briefly in turn.

(i) Deposit Insurance Pricing: Insurance premiums that are risk insensitive or only weakly sensitized to risk shift the burden of restraining risk-taking to the regulators rather than allowing risk to be controlled by the discipline of a pricing mechanism. This problem was recognized in FDICIA as we discuss below. There are, however, numerous difficulties in implementing a risk-sensitive deposit insurance pricing scheme that effectively deters risk-taking. These include risk measurement[7] and asymmetric information.[8] Properly calibrated risk-based deposit insurance pricing must depend on many variables, reflecting credit, interest rate, and liquidity risks. How do we measure these risks? Moreover, even if we could measure these risks, how should the deposit insurance premiums be linked to the measured risks? The asymmetric information problem arises from the bank having better information about its

7. See Kareken (1990) and Flannery (1991).
8. See Chan, Greenbaum, and Thakor (1992).

own risks than the regulator. Thus, in designing a deposit insurance scheme that accurately reflects risk, the regulator confronts the task of eliciting the bank's private information.

Risk-sensitive deposit insurance premiums were adopted under FDICIA, representing an important step in regulatory reform.

(ii) Regulatory Incentives: Some believe that many banking problems are rooted in defects in political and bureaucratic accountability.[9] Covering up evidence of poor regulatory performance and relaxing restrictions on regulated firms are common governmental responses to industry difficulties. Similarly, aggressive risk-taking by banks is a rational response to regulatory forbearance.

This viewpoint recognizes a principal-agent problem at the level of the public regulator. Regulators and politicians are seen as *agents* of taxpayers, and as agents they possess well-defined objectives that commonly conflict with those of their principals. To understand this viewpoint, imagine a banking or thrift industry that consists of many impaired firms with negative net worths that are attracted to risky portfolio strategies.

Regulators should expeditiously close such firms. But doing so usually upsets incumbent politicians. Moreover, resolute actions by regulators may signal their previous mistakes in allowing conditions to fester to the point where receivership or conservatorship becomes necessary. The larger the troubled firms' hidden economic losses, the more a public acknowledgement of their insolvency threatens the regulators' reputations, and the more inclined regulators will be to forbear. The hope may be that the insolvency can be reversed, or the problem can be passed on to a successor.[10] The idea that the careers of regulators and politicians would be damaged by acknowledging insolvencies is enshrined in the ancient practice of killing messengers bearing bad news. FIRREA and FDICIA dealt with this problem by greatly limiting regulatory forbearance; an institution whose book net worth falls below 2 percent of assets now has to be closed within a specified period of time.

If we assume that insured institutions constantly develop new and partly unanticipated ways to shift risk to the deposit insurance fund, we can imagine three *regulatory regimes* as shown in Figure 12.4.[11]

As shown in Figure 12.4, regulators start out as well-intentioned public servants. However, a "crossover point" is reached when they discover that actions taken by the firms they regulate have gone undetected and now threaten the solvency of the deposit insurance fund. The regulators recognize that the problems are so difficult to resolve that their career interests are better served by procrastination or denial. It appears that, in the case of the thrift industry, this transition to denial occurred in the late 1970s, and the denial continued through 1987. During this period, thrift-industry lobbying generated disinformation about the condition of the industry. Concern seemed limited to the possibility that public acknowledgement of the FSLIC's insolvency could precipitate a crisis. The existence of serious problems and regulatory denial is indicated by the data in Table 12.2.

9. The most forceful proponent of this argument is Kane (1989a, 1989b, 1990a, 1990b). See also Campbell, Chan, and Marino (1992) and Boot and Thakor (1993).

10. Boot (1992) has formalized this logic to show that managers may be tempted to hang on to negative NPV projects too long because divesting them may damage managerial reputation.

11. This is based on Kane (1989a, 1989b).

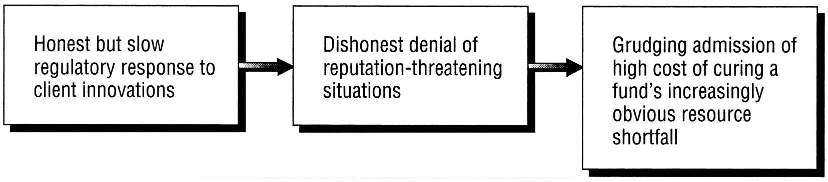


FIGURE 12.4 Sequential Regulatory Regimes

TABLE 12.2 Data on Financial Condition and Closures of FSLIC-Insured Thrifts from 1975 to 1984

Year	Ratio of Appraised Market Value of Net Worth to Total Assets in Percent	GAAP-Insolvent Institutions	Insolvencies Resolved by the FSLIC
1975	−7.77	17	11
1976	−7.25	48	12
1977	−6.62	38	10
1978	−6.87	38	4
1979	−9.32	34	4
1980	−12.78	43	32
1981	−15.41	85	82
1982	−10.63	237	247
1983	−6.03	293	70
1984	−2.74	445	336

Source: Edward J. Kane, "The Unending Deposit Insurance Mess," *Science* 246, October 1989b, 451–456.

As this table shows, the industry was insolvent on average for a long time before there was a public awareness. Moreover, closures or resolutions by the FSLIC fell far short of actual insolvencies. Since this was also a period during which the industry was growing rapidly, many firms with negative net worths were not only allowed to stay in business but also to grow their assets.

However, the longer a cover-up goes on, the harder it becomes to sustain the deception. The reasons are twofold. First, not acknowledging losses only defers accounting recognition. Second, allowing economically insolvent institutions to continue may increase losses over time given the bank's increased incentive to assume risk.

The above discussion takes as given the existing structure of federal deposit insurance. Many alternatives to this structure have been proposed, however. One is to *privatize deposit insurance*. This would provide a self-insurance program for the banking industry, in which banks insure and monitor each other. The difficulty is that, with a large number of banks, each member's incentive to monitor is likely to be weak.[12]

12. These incentives can be further weakened when some of the banks responsible for monitoring are in financial difficulty. The reduced monitoring by financially weak banks can lead to more banks getting into trouble due to the reduction of vigilance. This introduces a systematic risk element and possible contagions. For an analysis of the macroeconomic effects of bank runs, see Loewy (1991). See also Donaldson (1992).

A second alternative is to limit deposit insurance to "narrow banks." That is, only those banks that invest in the safest securities like Treasury bills would be able to offer insured deposit accounts. The remaining banks would invest in assets of their choice, but could finance these assets only with uninsured liabilities. This would limit taxpayers' exposure while preserving federal deposit insurance.

Finally, there are those who would reform deposit insurance by *eliminating the deposit contract itself*. As we saw in Chapter 10, the deposit contract creates the possibility of a bank run because of the *sequential service constraint* that entices each depositor to be the first at the teller window to withdraw his or her deposits as soon as trouble is suspected. While deposit insurance is one response to this potential disruption, another possibility is to do away with the sequential service constraint. This could be achieved if banks issued equity-like claims, such as those of mutual funds. Since any withdrawal from the fund would be commensurate with the fractional ownership of the investor in the fund, there would be no advantage to any investor/depositor in being first to withdraw. Such contracts could be endowed with the full range of transactions services. Individuals would be exposed to "market risk" since the value of the mutual fund would be subject to random fluctuations as market conditions change, but the threat of panic runs would vanish.[13] The issue was also discussed in Chapter 10. Figure 12.5 summarizes possible solutions to the deposit insurance problem.

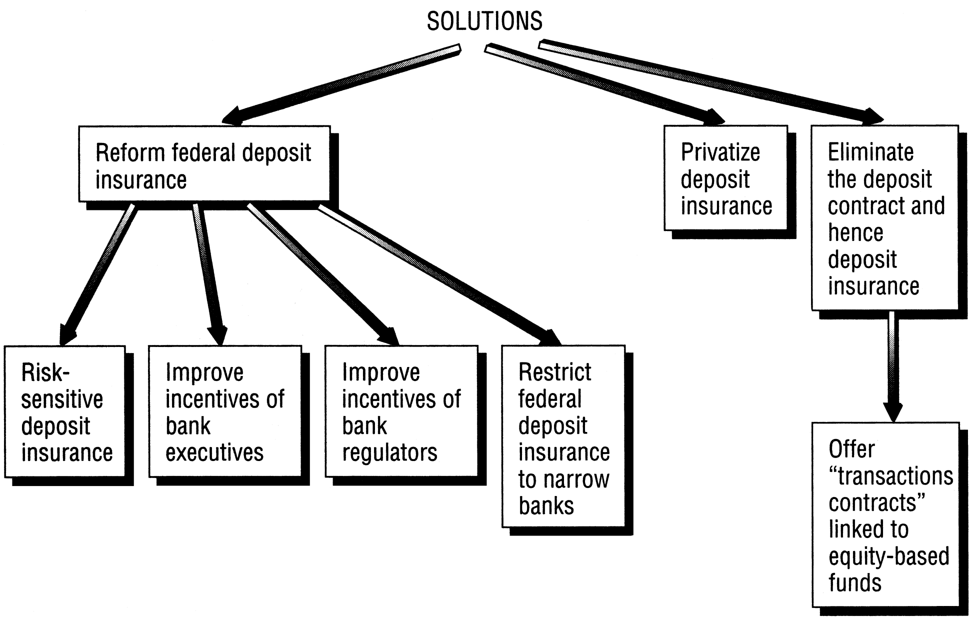


FIGURE 12.5 Possible Solutions to the Deposit Insurance Problem

13. Elimination of *non par clearance* was one of the motivations for the passage of the Federal Reserve Act of 1913. So to return to it may seem odd, but people are now willing to accept claims on mutual funds in lieu of deposits.

(iii) Improving the Incentives of Bank Executives: If the bank's top executives are somehow *rewarded* for failure, then the incentives to take risk are strengthened. One way that executives are rewarded is with *golden parachutes* when they leave troubled firms. The FDIC now has restrictions on such payments by troubled firms. This follows the OCC's decision in 1991 to stop monthly payments of $42,000 to Charles Zwick, former chairman of Southeast Banking Corporation. In February 1991, regulators forced Alan P. Hoblitzell, former chairman and CEO of troubled MNC Financial, Inc., to return $915,865 in severance pay.

The Bank Fraud Act of 1990 authorized the FDIC to set strict golden parachute rules in its role as guardian of the Bank Insurance Fund. The agency must approve any golden parachute offered to departing executives of troubled banks. To gain approval, the institution must demonstrate that the executive committed no fraudulent act, is not substantially responsible for the institution's impaired condition, and has not violated banking and criminal laws. Moreover, all institutions must satisfy similar criteria before they are allowed to pay legal fees for directors and officers.

- **Regulatory Uncertainty:** Public regulation of banking is aimed in part at dealing with moral hazard problems growing out of the public safety net. This includes not only explicit guarantees, such as deposit insurance and the LLR, but also ill-defined governmental guarantees of the payments system and TBTF. Thus, the question transcends that of designing an optimal regulatory monitoring system. Rather, it calls for *jointly* designing the guarantee structure *and* the monitoring system so as to achieve social objectives while controlling moral hazard at minimal *total* social cost.

An appealing approach is to minimize the *need* for regulation, but this means minimizing the span of the safety net. Indeed, this may be the most compelling argument for restricting the government's safety net. However, since even the minimal safety net will entail *some* government exposure, some regulation/supervision is likely to be necessary. We are, therefore, forced to wrestle with the question of optimal regulatory design in the context of a minimally guaranteed system.

In contemplating this issue, it is useful to distinguish between the *discretionary* and *nondiscretionary* aspects of regulation. The latter represent more or less well-defined rules, such as cash-asset reserve requirements, capital requirements, loans-to-one-borrower rules, and deposit insurance premium schedules. On the other hand, discretionary regulations involve greater *ambiguity*. Examples include the standards for bank holding company acquisitions contained in the 1970 Douglas Amendments, deposit insurance coverage under current practices, standards for access to the discount window, standards for intervention in cases of distressed institutions, and accounting standards in the banking and thrift industries.

One benefit of ambiguity is that it gives the regulator a weapon against moral hazard.[14] When a bank is not really sure whether the regulator will rescue it in a given set of circumstances, the bank may go to greater lengths to avoid jeopardy. However, ambiguity also has costs. As the probability and nature of regulatory intervention become more difficult to assess, investors begin to demand higher risk premia on the bank's equity. This increases the bank's cost of capital and reduces competitiveness relative to competitors in more predictable environments. To the extent that the

14. The constructive role of ambiguity/discretion has been stressed by Allen and Gale (1993), Boot and Thakor (1991), Boot, Greenbaum, and Thakor (1993), and Corrigan (1990).

regulator does *not* internalize the bank's increased cost of capital, discretion transfers wealth from shareholders to taxpayers or regulators.

• **Market Value Accounting:** It is widely believed that Regulatory Accounting Principles (RAP) and Generally Accepted Accounting Principles (GAAP) have contributed to recent problems in the thrift and banking industries. RAP hid the magnitude of the crisis for some time because it deferred unrealized losses and thereby overstated capital even when economic net worth was negative. Given the risk-taking incentives of economically insolvent institutions, the crisis gathered momentum as thrifts sought ever-increasing risk. GAAP do not help much since they rely substantially on the historical cost (or book value) of transactions.[15]

The alternative, *Market Value Accounting* (MVA), requires that all assets *and* liabilities, including all off-balance sheet items (which would be brought on-balance sheet) be carried at current market value. The values of the assets and liabilities would be increased or decreased, as market conditions indicated. MVA can be useful in implementing risk-based capital requirements, risk-based deposit insurance premiums, and improved regulatory supervision. It could, thus, be an important part of the overall reform of the deposit insurance system.

In principle, the case for adopting MVA is impeccable, but there are conceptual, measurement, and incentive problems with implementing MVA.[16] The major problem is assigning market values to *nontraded* commercial bank loans and guarantees. Many of these instruments are *nonmarketable* or marketable only at steep discounts. Because of information and monitoring advantages that the bank has relative to potential buyers of these instruments, there is usually a divergence between the value of the asset to the bank and its value if sold, so that measuring value becomes difficult.

Arguing that the system does not need a full-blown accounting system, but only a market-based measure of net worth, simplifies the measurement problem, but raises other issues that need further consideration. For example, it is possible that requiring banks to add capital when the market value of loans (and hence net worth) declines would induce banks to choose loans of shorter maturity, that is, loans that would "liquidate" with fewer possibilities for a market-based revision in value.[17] The reason is that there may be an asymmetry in the effect of revisions on the bank's capital. An increase in loan value would augment the bank's economic net worth and permit it to support deposit and asset expansion, but the bank may be unable to profitably carry out such an expansion immediately.[18] Thus, the bank may be unable to fully extract the benefit of an upward revision in loan value. But if the loan value drops, then MVA would force the bank to acquire additional capital or sell off some assets; both initiatives are likely to be costly.[19] Such an asymmetry in the effect of loan value

15. See White (1988).

16. See Berger, King and O'Brien (1991).

17. See O'Hara (1993) for an interesting model that predicts that, under some conditions involving asymmetric information, we could see a shortening of bank asset maturity with MVA, a veritable return to the concept of self-liquidating investments embodied in the "real bills" doctrine.

18. This would be the case, for example, if the marginal cost of deposits is increasing in quantity, and assets yielding more than the deposit funding cost are unavailable without sufficient time to plan for the expansion.

19. Again, this cost may be due to informational frictions. Asymmetric information alone can make equity capital more costly than deposits [see Myers and Majluf (1984)]. Moreover, loan sales may involve losses for the bank when it knows more than outsiders about the true value of the loans [see Bhattacharya and Thakor (1993)].

revisions on the bank could create an incentive to minimize potential value revisions, and hence a shortening of loan maturities. Despite these unresolved issues, valuing at market all those assets and liabilities that are actively traded in secondary markets and using best judgment to estimate the values of the remainder seems to be gaining favor among public regulators and the accounting profession.

• **Expanded Banking Powers:** Prior to the Gramm-Leach-Bliley Act of 1999, both banks and thrifts have lobbied for expanded powers that would increase charter values and reduce failures. The principal objections to *universal banks* were that expanded powers make it more difficult to regulate and limit risk-taking at the expense of the deposit insurance fund, and that expanded powers can give rise to conflicts of interest. The potential conflicts include:[20]

(i) The promotional role of the investment bank in selling securities may conflict with the commercial bank's obligation to provide objective advice to depositors.
(ii) In order to avoid a loss on a loan, the universal bank may encourage the borrower to raise new capital through the bank's securities subsidiary in order to repay the loan.
(iii) A universal bank may use its monopoly power to cross-sell services. Threats of credit rationing, refusal to renew loan commitments, and increasing the cost of loans could all be used to "tie" existing customers to other products of the universal banks.
(iv) A universal bank may avoid losses in underwriting by placing unsold securities in its trust accounts.
(v) Interlocks between the directors of universal banks and their customers may give rise to conflicts.
(vi) Banks may make imprudent loans to bolster a firm taken public through an *initial public offering* (IPO) underwritten by the universal bank.
(vii) The bank may lend imprudently to its securities affiliate, possibly transferring wealth from the deposit insurance fund to the securities affiliate.

The seriousness of these conflicts of interest is widely disputed.[21] Those favoring the separation of commercial and investment banking believe that conflicts are significant,[22] and that the separation is necessary to ensure that the governmental safety net is not significantly expanded in scope. Those favoring expanded bank empowerment believe that market forces will provide the discipline necessary to control abuses.[23] Others believe that moral hazard and other conflicts of interest can be controlled by limiting regulatory forbearance and forcing sufficiently early closure of troubled institutions,[24] as has now been done.

If there are serious conflicts, then there should be evidence that banks in the pre-1933 period deceived investors into investing in securities that imposed losses on

20. See Saunders (1985).
21. See Benston (1990).
22. See Mester (1992).
23. See Huertas (1988) and Saunders (1991). For historical accounts of the separation of investment banking, see Shull (1983).
24. See Eisenbeis and Horvitz (1993).

investors. A recent study tested this hypothesis by comparing the performance of securities underwritten by affiliates of commercial banks with those sponsored by independent investment banks.[25] It found that securities sold by bank affiliates defaulted less frequently than similar securities sold by stand-alone investment banks. Strikingly, the difference in default rates was the greatest for relatively speculative (private-information-intensive) issues that, because they are the hardest for investors to judge, should have potentially imposed the largest losses on investors. This evidence calls into question the significance of alleged conflicts of interest. Investors with rational expectations will take into account potential conflicts of interest in pricing securities. Consequently, securities sold by bank affiliates with poor reputations for avoiding potential conflicts will sell at steep discounts. Issuers will anticipate this and gravitate to bank affiliates with good reputations, and these institutions will then be observed to underwrite the majority of issues accounted for by bank affiliates. Succinctly put, market discipline apparently worked well in resolving conflicts of interest prior to Glass-Steagall.

An additional argument against expanded banking powers is that they may dilute financial innovation incentives. A universal bank that is considering a financial-market innovation will worry about cannibalizing the loan business of its commercial banking arm. A stand-alone investment bank has no such concerns and thus will have stronger financial-innovation incentives.[26]

The effect of expanded powers on banks is ultimately an empirical issue. Barth, Caprio and Levine (2004) provide interesting evidence. Using data on 107 countries, they show that restricting bank activities has an adverse effect on bank development and stability. This provides strong support for universal banking, which is common in Europe, Latin America and other parts of the world.[27]

The 1991 FDICIA and Beyond

In 1991, the U.S. Treasury proposed sweeping regulatory reform of banking aimed at promoting the global competitiveness of American banking institutions, reducing taxpayers' exposure deriving from deposit insurance, and promoting the safety and soundness of American financial institutions. The key elements of the proposal were: (i) limiting deposit insurance coverage, (ii) achieving regulatory consolidation, (iii) involving the Treasury as well as the Federal Reserve in TBTF decisions, and (iv) dismantling the Glass-Steagall and McFadden restrictions on banking activities.

This initiative culminated in the Federal Deposit Insurance Corporation Improvement Act (FDICIA) in December 1991. The main focus of FDICIA was on reducing taxpayer exposure deriving from deposit insurance and promoting the safety and soundness of American financial institutions. The key features of FDICIA are discussed below.

Bank Regulation FDICIA linked *supervision* to *bank capital*. Regulators were required to establish five capital compliance categories for banks and thrifts: well-

25. See Kroszner and Rajan (update).

26. See Boot and Thakur (1997) who suggest that this may be why investment banking is more advanced in the U.S. than in Europe.

27. See Mester (1992).

capitalized, adequately capitalized, undercapitalized, significantly undercapitalized, and critically undercapitalized. Regulatory forbearance was restricted by requiring "prompt corrective action" as capital dissipated. In particular, regulators are required to close banks *before* they become insolvent. If capital declines to levels below positive trigger points, regulators must impose caps on growth, enforce reductions or suspension of dividends, instruct bank management to raise capital, and mandate management changes if necessary. Regulators are also permitted to close critically undercapitalized banks, where the ratio of tangible equity capital to total assets is less than 2 percent. FDICIA also permits bank regulators to place a bank in receivership or conservatorship for other transgressions, including violation of a cease-and-desist order, concealment of records or assets, inability to cover deposit withdrawals, or failure to either develop or implement a required plan to raise capital. Moreover, FDICIA requires bank regulators to take action within 90 days of a bank becoming critically undercapitalized.

Prompt corrective action also requires an ex post review of any bank or thrift failure that imposes material costs on the FDIC. If a material loss occurs, the inspector general of the appropriate banking agency must determine why and must provide recommendations for preventing such a loss in the future. This report must be made available to the Comptroller General of the United States, to members of Congress upon request, and to the public through the Freedom of Information Act. Further, the GAO must do an annual review of the reports and recommend improvements in supervision.

In addition to supervision, regulators were instructed to come up with a way to link bank capital requirements to interest-rate risk, credit risk of concentrations of credit, and the risk of nontraditional activities, and to draft a new set of noncapital measures of bank safety, such as underwriting standards. Regulators also are required to perform annual on-site bank examinations, place limits on real estate lending by banks and tighten auditing requirements.

Deposit Insurance FDICIA transferred to the FDIC the responsibility for insuring thrifts as well as commercial banks. (Credit unions continue to have a separate deposit insurance agency, the NCUA.) While the deposit insurance of thrifts and banks was consolidated into the FDIC, the two types of intermediaries retained separate insurance reserve funds—bank deposits are insured by the *Bank Insurance Fund* (BIF) and thrift deposits by the *Savings Association Insurance Fund* (SAIF).

FDICIA also reduced the scope of federal deposit insurance. The most significant change is the restriction on the TBTF initiatives that provide governmental protection of deposits beyond the prescribed $100,000 limit. The FDIC's ability to reimburse uninsured depositors—those with over $100,000 and those with foreign deposits—was severely limited. FDICIA, however, does permit TBTF initiatives if failure would "have serious adverse effects on economic conditions or financial stability." This exception requires the agreement of a two-thirds majority of the directors of the FDIC, a two-thirds majority of the Board of Governors of the Federal Reserve System, and concurrence of the Secretary of the Treasury. The Secretary of the Treasury is required to document the need to invoke the systemic risk exception. The GAO must review any actions taken, and analyze the potential effect on the behavior of other insured depository institutions as well as uninsured depositors. The rest of the banking industry is required to pay the cost of any bailout through an emergency assessment by the FDIC that is proportional to each bank's average total tangible assets.[28] Only the best-capitalized banks will be able to offer *insured brokered deposits* (large CDs sold through brokerage firms) or accounts established under employee pension plans that offer pass-through insurance.[29] FDICIA also required the FDIC to adopt *risk-sensitive deposit insurance premia.*

FDIC Funding The shrinking bank deposit insurance fund was bolstered with an additional \$70 billion in borrowing authority. The FDIC's authority to borrow from the Treasury was increased from \$5 billion to \$30 billion. The loans were to be repaid with increased deposit insurance charges on the banks. The FDIC was authorized to borrow additionally for working capital needs. The money, about \$45 billion, and interest would be repaid as the FDIC gradually disposes of the assets of failed banks. The FDIC also was instructed to rebuild the BIF to 1.25 percent of insured domestic deposits by the year 2006, which was achieved well before that.

The Discount Window FDICIA limits the Federal Reserve's ability to use the discount window to support a financially troubled bank. Permitting discount window access to a failing bank allows uninsured deposits to be withdrawn prior to FDIC resolution, thereby increasing the exposure of the deposit insurance fund. FDICIA limits the amount of discount window lending to a bank's capital, with restrictions applying to undercapitalized (capital less than 8 percent of assets) and critically undercapitalized banks. Although the Federal Reserve retains considerable discretion in its discount window policy, it is liable to the FDIC for losses suffered by the deposit insurance fund due to discount window access provided to critically undercapitalized banks.

Another significant change in the discount window is that FDICIA now permits *all* nonbank firms—brokerage and other financial services firms as well as nonfinancial firms—to borrow at the discount window for emergency purposes under the same collateral terms afforded to banks.

Corporate Governance of Banks FDICIA contains provisions aimed at strengthening the audit function of the boards of directors of banks and developing guidelines for the compensation of directors and officers. These provisions are designed to protect the deposit insurance fund by enhancing managerial and director accountability.

Specifically, banks are required to have audit committees composed exclusively of "outside" directors who are independent of the management of the institution. Two additional requirements are imposed on large institutions. First, their audit committees cannot include large customers of the institution. Second, audit committee members must have banking or related financial management expertise, and they must have access to independent outside counsel of their own choosing. FDICIA prescribes that the audit committee shall review external audits with management and

28. FDICIA thereby introduces an incentive for insured banks themselves to question any TBTF initiative.

29. Pass-through insurance refers to \$100,000 coverage for each participant in a collective account. Thus, a \$50 million pension fund deposit might be fully insured, provided it had a sufficient number of participants $\left(\frac{\$50\text{ million}}{\$100.00} = 500\text{ participants}\right)$.

the independent accountants. These provisions are designed to increase the independence of the audit committee and its ability to monitor management. They in many ways anticipate the requirements of the Sarbanes–Oxley legislation of 2001 which applies to all publicly owned corporations.

FDICIA's impact on board compensation committees is less direct. FDICIA does not specify the composition of the board's compensation committee, but it calls for federal banking agencies to prescribe guidelines for executive and board compensation that preclude employment contracts that could jeopardize the financial health of the institution.

Foreign Banks and Foreign Deposits FDICIA gives the Federal Reserve new authority to regulate foreign bank operations in the United States. The FDIC is generally prohibited from protecting foreign branch deposits of a failed bank. In cases where the agency determines offshore deposits must be repaid to protect the system, it is required to recover losses through an industry-wide assessment on an expanded base that has the effect of assessing foreign deposits.

Accounting Reforms The federal banking agencies must issue regulations requiring banks to report *off-balance-sheet items* on financial statements. In addition, the agencies must require disclosure of the fair market value of all assets, to the extent possible.

Restrictions on State Bank Powers FDICIA prohibits state banks from exercising powers not permissible to federally chartered institutions, including insurance underwriting. The bill "grandfathers" banks already lawfully engaged in underwriting insurance under state law. Another exemption permits state banks to invest up to 10 percent of their portfolio in stocks listed on national securities exchanges, provided they are already in the business.

Consumer Provisions The principal consumer protection is the truth-in-savings provision, which requires uniform disclosure of the terms and conditions of savings accounts. A "greenlining" amendment provides incentives for banks to lend money in less affluent neighborhoods. The FDIC is required to start an affordable housing program and to give nonprofit organizations an opportunity to purchase residential properties acquired from failed banks.

Miscellaneous Provisions FDICIA also relaxes the "qualified thrift lender test" allowing thrifts to invest more of their assets outside of housing-related areas.

An Evaluation of FDICIA[30] As we have seen in earlier chapters, banks must process risk if they are to serve as qualitative asset transformers. However, because of the regulatory safety net that is needed to foster banking stability, a moral hazard arises stemming from banks' propensities to take *excessive* risks. The goal of bank regulation should be to address this moral hazard without stifling the intermediation function of banks.

30. For additional readings on FDICIA, see Booth (1993), Carnell (1992), Greenspan (1993), and Wall (1993).

FDICIA focused on limiting the deposit insurance exposure of taxpayers.[31] To this end, FDICIA provides banking agencies with a clear goal of minimizing deposit insurance losses and providing incentives to encourage compliance. The prompt corrective action requirements limit regulatory forbearance, and risk-sensitive capital requirements and deposit insurance premia may encourage risk abatement.[32]

There are four main criticisms of FDICIA. First, while the principal thrust of FDICIA was to limit the size and scope of the federal financial safety net, the discount window access given to nonbanking firms potentially expands the safety net.[33]

Second, FDICIA may impede desirable risk-taking function of banks as qualitative asset transformers. FDICIA was a reaction to perceived excesses in the industry and to failures of the regulators. However, FDICIA failed to address the question of *optimal* risk-taking by banks[34] Rather, it focused on recapitalizing the BIF and ensuring that future costs to the deposit insurance fund were better controlled. But this could deter banks from processing the kinds of risks that are also socially optimal.

Third, FDICIA directs each federal banking agency to monitor banks' operations, management, asset quality, earnings, stock values, and the compensation of executives and directors. This mandate could pressure regulators to "micromanage" banks and to discourage desirable risk-taking and innovation.[35]

Fourth, FDICIA failed to address the issue of the competitiveness of U.S. banks *vis à vis* foreign banks and nonbank competitors. The discretionary elements of regulation, including expanded powers to fine and dismiss directors and officers and to review executive compensation, raises potentially nondiversifiable investor risk and, therefore, increases banks' cost of capital. This reduces banks' competitiveness. Also, the increased exposure of directors and officers elevates the cost and difficulty of staffing these positions. While the Act will require greater care on the part of directors and officers, it will discourage those with reputational capital or other forms of wealth from serving as directors and officers of financial institutions.

Another competitive weakness of FDICIA relates to Glass-Steagall and interstate branching prohibitions. When the 1991 Treasury proposal was being discussed, the financial health of banks was thought to be too precarious to permit expanded powers via repeal of Glass-Steagall. When the issue resurfaced in 1993, opponents of repeal argued that bank profits were at historic highs, so that expanded powers were unnecessary to bolster bank profitability. The key impediment to repeal of

31. The focus of FDICIA was not a new concern as the following 1933 rhetorical question by Senator Carter Glass indicates.

"Is there any reason why the American people should be taxed to guarantee the debts of banks, any more than they should be taxed to guarantee the debts of other institutions, including the merchants, the industries, and the mills of the country?"

See Smith and Beasley (1972).

32. The crudity of the risk sensitizatons may encourage new circumventions with attendant deadweight losses. Some also criticize FDICIA for not going far enough in restricting regulatory forbearance. Many of the important changes are suggestions, so that regulatory discretion remains in the treatment of problem banks. See Carnell (1992). Discretion, however, can be of value to regulators in dealing with banks in a rapidly changing environment. See Boot, Greenbaum, and Thakor (1993).

33. Whether this represents a significant risk exposure depends on how easily the Federal Reserve grants access to nonbanks.

34. Greenspan (1993) noted, "The legislative and regulatory process, in my judgment, has never adequately wrestled with the question of just how much risk is optimal."

35. See Greenspan (1993).

Glass-Steagall appeared to be the potential for increased taxpayer exposure. Further, the greater number of regulatory requirements (including requirements such as Truth-in-Savings that are unrelated to bank safety) could further damage banks' competitiveness.

The Financial Services Modernization Act of 1999

After a relentless weakening of its key separation provisions, the Glass-Steagall Act was finally formally dismantled in 1999 with the passage of the Financial Services Modernization Act or the Gramm-Leach-Bliley Act. This act repealed Sections 20 and 32 of the Glass-Steagall Act. It also authorized bank holding companies and foreign banks that meet eligibility criteria to become financial holding companies, thus allowing them to engage in a broad array of financially related activities. In addition, the Act addressed the functional regulation of financial holding companies, the protection of nonpublic customer information held by financial institutions, the supervision of the CRA, and other regulatory practices. A summary of the key elements of the act are provided below.

- **Repeal of Glass-Steagall**. Allows U.S. financial services providers, including banks, securities firms and insurance companies to affiliate with each other and enter each other's markets.
- **Bank Holding Company Structure**. Generally bank holding company affiliates will be the vehicles through which to engage in a broad range of financial activities.
- **Qualification to Engage in Financial Activities**. Requires all subsidiary insured depository institutions of the holding company to be well capitalized and well managed in order for the holding company to engage in broader financial activities. Divestiture and/or other restrictions and limitations may be required in the event of noncompliance.
- **Operating Subsidiary Activities**. Allows national banks with assets of $1 billion or less to conduct financial activities through operating subsidiaries. In order to conduct such activities through a subsidiary, the national bank and all insured depository institution affiliates must be well capitalized and well managed and the national bank must receive the approval of the OCC based on those criteria. A national bank subsidiary engaging in such activities will be subject to affiliate transaction restrictions and to antitying prohibitions. The bank also must deduct from capital the amount of its investment in the subsidiary. National banks with assets exceeding $1 billion must conduct financial activities through holding company affiliates. National banks of any size may engage in financial activities on an agency basis through an operating subsidiary. National banks lawfully conducting activities through operating subsidiaries as of the date of enactment will be permitted to continue such activities.
- **Municipal Revenue Bond Underwriting**. Authorized as a permissible banking activity. Therefore, this activity may be conducted by the bank directly or in an operating subsidiary. Previously, only the general obligations of municipalities could be underwritten by banks.

- **Functional Regulation**. Relies on strong functional regulation of the banking, insurance and securities components of the holding company, and establishes the Federal Reserve as the umbrella regulator.
- **Reduces Regulatory Burdens**. Streamlines regulatory burdens by requiring the Federal Reserve as umbrella supervisor to rely on reports and examinations conducted by other functional regulators. Also requires sharing of information among affected regulatory agencies as necessary to carry out their official duties.
- **Competition Protection Rules**. Requires the federal banking agencies to issue joint consumer protection regulations governing the sale of insurance products by banks, their employees, or others who engage in such activities on behalf of the banks. The federal banking regulators must consult with the states in the process of formulating their joint rules. Provisions of federal rules deemed more protective will pre-empt state law or rules unless within three years of federal notification the state legislatures enact laws opting out of such coverage.
- **FICO Assessment**. Freezes the BIF-member FICO assessment for three years, beginning 1999. This represents a saving of about $18,000 per year for a bank with total assets of $100 million. It is therefore important to smaller community banks. This freeze is important because it will give Congress time to consider other important issues such as the merger of the FDIC insurance funds, merger of banks and thrift charters, and consolidation of regulatory agencies such as the OCC and the OTS. Since 1980, there have been at least 13 congressional hearings on the soundness of the federal deposit insurance system. There have been 11 proposals introduced concerning consolidation of federal regulation of banks and savings and loan institutions; and five proposals to merge bank and thrift charters.
- **CRA**. Establishes a rebuttable presumption of CRA compliance with respect to an insured depository institution that has achieved a "satisfactory" or better rating in its most recent CRA exam and in each of its CRA exams during the immediately preceding 36-month period. The presumption of compliance may be rebutted by any person presenting substantial verifiable information to the contrary.

  Banks and savings and loan associations with total assets less than $100 million and located in nonmetropolitan areas are exempted from the provisions of the CRA. This exemption only applies to 38 percent of all banks and savings and loans, which collectively control only 2.8 percent of banking assets nationwide.

- **Bank Securities Activities**. While eliminating the broad exemption that banks enjoyed from registration as a broker or dealer under the securities laws, the bill makes clear that banks serving as custodians to self-directed IRAs will not be required to push these activities out of the bank and into a registered broker or dealer. Banks often function as service providers to pension, retirement, profit sharing, bonus, thrift, savings, incentive and other plans. The SEC, with the concurrence of the Federal Reserve Board, may determine by regulation those new products which, if offered or sold by a bank, would subject it to registration with the SEC. A bank may offer or sell "traditional banking products," as defined in this section, without becoming subject to registration with the SEC.
- **Federal Home Loan Bank Reforms**. Includes provisions to modernize the operations of the Federal Home Loan Bank System. As of June 1, 2000, membership

in the Federal Home Loan Bank System was made voluntary. Community banks (those banks with total assets less than \$500 million) will be able to become members without regard to the percentage of total assets represented by residential mortgage loans. Community banks will be able to use advances for small business, small farm and small agribusiness lending. Also allows community banks to collateralize advances with small business and agricultural loans. Modifies the governance structure of the System to give more authority to the regional banks.

Liquidity Constraints, Capital Requirements, and Monetary Policy

The size and composition of banks' balance sheets are constrained by the legal reserve requirement, which establishes a minimum ratio of cash assets to deposit liabilities, and the capital requirement that establishes minimum ratios of bank capital to risk assets (loans, to a first approximation) and to total assets (the so-called leverage ratio). For most of U.S. history, and certainly for the half-century following the Great Depression, capital requirements tended to be without effect or not binding. Therefore, the operating constraint on banks' size was the legal cash-asset reserve requirement.[36] This changed in the late 1980s when increased credit risks of banks' on- and off-balance-sheet activities were recognized. At the same time, legal cash-asset reserve requirements were reduced in a series of steps, at least partly in response to the dismal record of bank earnings. For many banks, the position of reserve requirements and capital requirements was reversed, so that the capital requirement became binding. This affected the way monetary policy, especially the Federal Reserve's open-market operations, played out in the economy.

In the traditional setting, Federal Reserve purchases of government securities would expand the *excess* cash-asset reserves of the banking system, prompting the banks to expand lending and the asset size of their balance sheets. The initial Federal Reserve purchase would drive up U.S. government securities prices and depress interest rates. The secondary effect of the expansionary open-market operations would come from the banks' reactions to new deposits from those who sold their government securities to the Federal Reserve. These autonomous deposits increase the banks' cash-asset reserves. In an effort to dissipate the new excess reserves, the banks lend newly created deposits. This expands the asset (liability) size of the banks and exerts added downward pressure on interest rates.

All of this would take place without any interference or influence from capital requirements. Indeed, this traditional interpretation of monetary policy assumes that the capital requirement is not binding. If, however, the capital requirement is binding, a very different picture emerges. The initial effects on interest rates via the government securities markets remains unchanged, but bank reactions are conditioned by an altered constraint. To be sure, a capital constraint can be relaxed by reducing dividends or selling bank equity, or even by realizing capital gains by selling assets or liabilities that are being carried at understated historical values. But such adaptations often are costly in the short run. Thus, the bank's capital may be fixed in the short run

36. Even today, most money and banking textbooks explain the deposit expansion process as if the capital requirement is without effect. See our discussion of the fixed coefficient model in Chapter 3.

as a practical matter, and in such cases, it is the other balance sheet accounts that must accommodate to the capital constraint.

Capital requirements are of two types: risk-based and leverage ratios. For simplicity, we can think of the former as a minimum ratio of capital to loans and the latter as a minimum ratio of capital to total assets (liabilities). Now suppose the bank can hold nonearning cash assets, credit risk-free, interest-bearing government securities, or risky loans. Assume further that the cash-asset reserve requirement is zero, but there is a binding capital requirement of either the risk-based or leverage type. Banks have one class of noninterest bearing deposits and one type of capital, equity.

If the capital requirement is of the risk-based variety, and the Federal Reserve undertakes an expansionary open-market operation, the bank will receive an autonomous deposit. Some small fraction of the deposit inflow may be held in the form of cash for liquidity purposes, but most will flow into government securities. We know the bank will *not* make loans because it is capital constrained. Thus, interest rates will fall, deposit expansion will ensue, *but loans will not be made*.

Binding capital requirements, therefore, can explain a *credit crunch* even though monetary policy retains its effectiveness in terms of influencing interest rates. Interest rates are lowered by the Fed's expansionary open-market operations, first as a result of the Federal Reserve's purchase of government securities, and then as a result of the bank's purchase of government securities. The drop in interest rates will presumably spur investors to borrow via the capital markets, but not all have easy access to these markets. Thus, the expansionary impact of a given open-market operation may well be weakened, and it will certainly be rechanneled with smaller firms more likely to be stifled.

Now consider a leverage-type capital ratio that fixes the maximum amount of total assets the bank can hold for any given amount of capital. In this case, the banks are totally out of the monetary policy loop. A purchase of government securities by the Federal Reserve cannot produce even an initial increase in bank deposits. Since the bank's size is constrained by the capital requirement, it cannot accept the deposit of the seller of government securities unless it simultaneously eliminates another deposit of equal amount. In this interesting case, the initial purchase of government securities by the Federal Reserve will put expansionary downward pressure on interest rates, but the banks will not be able to expand. An autonomous deposit will require the bank to sell loans or securities in order to extinguish an equivalent amount of deposits. This will put upward pressure on interest rates and force banks to hold excess reserves.[37]

Supplanting cash-asset reserve requirements with capital requirements inevitably alters the way in which monetary policy affects the economy at large, and this has special relevance for understanding the 1991–92 *credit crunch*. Bank loans were said to be unavailable despite the Federal Reserve's efforts to stimulate the economy by lowering short-term interest rates.[38] The Federal Reserve actively purchased government securities, expanding the reserves of the banking system, and bank assets grew, but lending remained largely unaffected. Banks simply increased their holding of government securities, and the economy grew sluggishly until the end of 1993.

37. To firm your understanding of this analysis, consider the possibility of having the Federal Reserve do its open-market operations by buying and selling bank equity instead of U.S. government securities.

38. Most Federal Reserve open-market operations are in the short end of the government securities market. This is because the short end of the market is deeper and more liquid and, therefore, can accept the Federal Reserve's large transactions with relatively little disruption.

The Basel II Capital Accord

In June 2004, central bank governors and heads of bank supervisory authorities in the Group of Ten (G10) countries issued a press release and endorsed the publication of **International Convergence of Capital Management and Capital Standards: A Revised Framework**, a new capital adequacy framework, commonly known as Basel II. The planned implementation of the basic approach was to be completed by end of 2006, with the more advanced approaches to be adopted by end of 2007. Basel II is viewed by many as the outcome of a process of evolution started by Basel I.

Objectives of Basel II

Basel II has numerous objectives. The main ones are listed below:

- Ensure that capital adequacy regulation is not a source of competitive disadvantage.
- Adopt more risk-sensitive capital requirements.
- Make greater use of banks' own internal risk assessments.
- Bring market discipline and regulatory monitoring to bear as part of regulation to ensure prudent risk-taking rather than relying solely on capital requirements.
- Cover a more comprehensive set of risks, including credit risk, interest rate risk and operational risk.
- Account for the risk mitigation efforts of banks.
- Adopt a more forward-looking approach that can evolve with time.

The Three Pillars of Basel II

Basel I focused exclusively on bank capital requirements. In contrast, Basel II takes a more comprehensive approach, relying on three "pillars" to ensure appropriate risk-taking by banks. These three pillars are:

- **First pillar:** Minimum Capital Requirements;
- **Second pillar:** Supervisory Review Process; and
- **Third pillar:** Market Discipline.

The idea is that regulators are supposed to rely on three mechanisms for controlling bank risk: capital requirements (as in Basel I, but with modifications to link capital requirements to a broader array of risks than just credit risk), regulatory monitoring, and market discipline. We discuss each pillar briefly in what follows. The interested reader should visit the Bank for International Settlements Web site for a more detailed discussion.

The First Pillar: Minimum Capital Requirements

Total minimum capital requirements have to be calculated for credit, market and operational risk. The capital ratio is calculated using the definitions of regulatory capital and risk-weighted assets. The total capital ratio must be no lower than 8 percent. Tier-2 capital is limited to 100 percent of Tier-1 capital.

Definition of Eligible Regulatory Capital: With a few modifications, this is essentially the same as in the 1988 Basel I Accord.[39]

Definition of Risk-Weighted Assets: Basel II defines

Total Risk-Weighted Assets = [Risk-weighted assets determined by credit risk] +[12.5 × Capital requirement for market and operational risks].

Note that the 12.5 above is the reciprocal of the minimum capital requirement of 8 percent.

Determination of Risk-Weighted Assets for Credit Risk

Banks are allowed a choice between two broad methodologies for calculating their capital requirements for credit risk: the *standardized approach* and the *internal ratings-based* (IRB) approach. We consider the **standardized approach** first. In Table 12.3 below, we provide the weights to be assigned for different kinds of credits under this approach.

In addition to stipulating risk weights to reflect credit risks embedded in a variety of different assets, Basel II also recognizes that the *risk mitigation* efforts of banks can affect their risk exposure, and seeks to account for this in the computation of minimum capital requirements. For example, the bank may be able to lower the capital it posts against a transaction if it is *collateralized*, i.e., the bank's credit exposure is limited by collateral. If, however, the claim in question has an issue-specific rating that reflects the bank's risk mitigation efforts, then no additional capital reduction is granted beyond what is already made possible by the effect of the risk mitigation on the credit rating.

Next, we turn to the **IRB** approach to credit risk. This approach permits some banks to rely on their own estimates of risk components in determining the capital requirement for a given exposure, as long as the banks using this approach meet certain conditions and disclosure requirements. The risk components in the IRB approach include measures of the probability of default (PD), loss given default (LGD), the exposure at default (EAD), and effective maturity (M). In some cases, banks may be required to use a supervisory value as opposed to an internal estimate for one or more of the risk components.

The first step in the IRB approach is to categorize banking-book exposures into broad classes of assets with different underlying risk characteristics: corporate, sovereign, bank, retail, and equity. Within those broad classes, there are sub-categories.

For each of these asset classes, there are three key elements:

- **Risk Components:** estimates of risk parameters provided by banks, some of which are supervisory estimates;
- **Risk-Weighted Functions:** the means by which risk components are transformed into risk-weighted assets and therefore capital requirements;
- **Minimum Requirements:** the minimum standards that must be met in order for a bank to use the IRB approach.

39. For example, under one of the permissible approaches (the internal ratings-based approach), the treatment of including general loan-loss reserves in Tier-2 capital is withdrawn.

TABLE 12.3 Risk Weights for Different Credits Under Standardized Approach

Types of Claims	Risk Weights Assigned
1. Claims on sovereign governments and their central banks	Depends on credit ratings: 0% for AAA to AA−; 20% for A+ to A−; 50% for BBB+ to BBB−; 100% for BB+ to B−; 150% below B−; and 100% if unrated.
2. Claims on noncentral government public sector entities (PSEs)	Risk-weighted at national discretion, with claims on certain domestic PSEs being treated as claims on the sovereigns in whose jurisdictions the PSEs are established.
3. Claims on multilateral development banks (MDBs)	Risk weights are based on external risk assessments, with a 0% risk weight applied to claims on highly rated MDBs (e.g. those with external assessments of AAA).
4. Claims on banks	National supervisors can choose from one of two options: (i) assign all banks incorporated in a given country a risk weight one category less favorable than that assigned to claims on the sovereign of that country, with a cap of 100% on the risk weight; or (ii) base the risk weighting on the external credit assessment of the bank itself, subject to a floor of 20% and claims on unrated banks being risk weighted at 50%.
5. Claims on securities firms	To be treated as claims on banks if securities firms are subject to supervisory and regulatory arrangements similar to banks, including risk-based capital requirements; otherwise, the rules for claims on corporates apply.
6. Claims on corporates	Depends on credit ratings: 20% for AAA to AA−; 50% for A+ to A−; 100% for BBB+ to BB−; 150% below BB−; and 100% for unrated. At national discretion, supervisory authorities may permit banks to risk weight all corporate claims at 100% without regard to external ratings.
7. Claims included in regulatory retail portfolios, such as revolving credit and lines of credit (such as credit cards and overdrafts), personal term loans and leases, and small-business facilities and commitments.	Risk-weighted at 75%
8. Residential mortgages and claims secured by residential property.	Risk-weighted at 35%
9. Claims secured by commercial real estate.	Risk-weighted at 100%
10. Past due loans (past due for more than 90 days).	Risk-weighted at 100% to 150% depending on specific provisions.
11. High-risk categories such as claims on sovereigns, PSEs, banks, and securities firms rated below B−, claims on corporates related below BB−, securitization tranches rated between BB+ and BB−.	Risk-weighted at 150% or higher, with securitization tranches rated between BB+ and BB− risk-weighted at 350%
12. Other assets like investments in equity or regulatory capital instruments issued by banks or securities firms.	Risk-weighted at 100%
13. Off-balance sheet items	• Credit Conversion Factors (CCFs) will be used. Commitments with an original maturity of up to 1 year will receive a CCF of 20%, commitments of original maturity over 1 year will receive a CCF of 50%, whereas commitments with a Material Adverse Change (MAC) clause receiving a 0% CCF. • Short-term self-liquidating trade letters of credit will receive a 20% CCF.

For many of the asset classes, there are two broad approaches: a **foundation approach** and an **advanced approach**. Under the foundation approach, as a general rule, banks provide their own estimates of PD and rely on supervisory estimates for other risk components. Under the advanced approach, banks provide more of their own estimates of PG, LGD, and EAD, and their own calculation of M. We skip the details of these calculations here.

Securitization receives special treatment under Basel II, and banks are required to determine regulatory capital requirements on exposures arising from traditional and synthetic securitizations, keeping in mind the economic substance of the securitization rather than its legal form. The securitization structures subject to capital requirements include exposures arising from the provision of credit-risk mitigants to a securitization transaction, investments in asset-backed securities, retention of a subordinated tranche, and extension of a liquidity facility or credit enhancement. The actual capital requirements against these exposures depend on the credit ratings of the exposures.

Capital Requirements Against Operational Risk: Operational risk is defined as the risk of loss resulting from inadequate or failed internal processes, people or systems or from external events. It includes legal risk, but excludes strategic and reputational risk.

There are methods for calculating operational risk capital charges: (i) the Basic Indicator Approach; (ii) the Standardized Approach; and (iii) Advanced Measurement Approaches (AMA). Banks are encouraged to move along the continuum of available approaches as they develop more sophisticated operational risk measurement systems and practices, with specific qualifying criteria specified for the Standardized Approach and the AMA.

Under the **Basic Indicator Approach**, the bank must hold capital for operational risk equal to 15 percent of positive average annual gross income for the previous three years. Figures for any year in which annual gross income is negative are excluded.

Under the **Standardized Approach**, banks' activities are divided into eight business lines: corporate finance, trading and sales, retail banking, commercial banking, payment and settlement, agency services, asset management, and retail brokerage. The total capital charge is calculated as the three-year average of the simple sum of the regulatory capital charges across each of the business lines in each year. In any given year, negative capital charges (resulting from negative gross income) in any business line may offset positive capital charges in other business lines without limit. However, when the aggregate capital charge across all business lines within a given year is negative, then that year is excluded from the calculations. The percentages of gross income to be kept as capital vary across business lines.[40]

Under the **AMA**, the regulatory capital requirement equals the risk measure generated by the bank's internal operational risk measurement system. Use of the AMA is subject to supervisory approval.

Capital Requirements Against Market Risk: Under Basel II, banks are required to have procedures that enable them to assess and actively manage all material market risks at position, desk, business line, or firm-wide level. The assessment of internal capital adequacy for market risk should be based on both **Value-at-Risk (VAR)**

40. It is 18 percent for Corporate Finance, Sales and Trading, and Payment and Settlement; 15 percent for Commercial Banking, and Agency Services; and 12 percent for Retail Banking, Asset Management, and Retail Brokerage.

modeling and stress testing, including an assessment of concentration risk and assessment of illiquidity under stressful market scenarios. The bank's internal capital assessment is required to demonstrate that it has enough capital to not only meet the minimum capital requirements but also to withstand a range of severe but plausible market shocks.

Wherever appropriate, banks are required to factor in:

- illiquidity/"gapping" of prices;
- position concentration (relative to market turnover);
- nonlinear products/deep out-of-the-money positions;
- events and jumps-to-defaults;
- significant shifts in correlations;
- other risks that may not be appropriately captured by VAR, such as recovery rate uncertainty, and skewness risk.

The Second Pillar: Supervisory Review Process

The supervisory review process of Basel II is intended to ensure that banks have adequate capital to support all the risks in their business, but also to encourage banks to develop and use better risk management *techniques* in monitoring and managing their risks. This review process recognizes the responsibility of bank management in developing an internal capital assessment process and setting appropriate capital targets.

Supervisors are expected to evaluate how well banks are assessing their capital needs relative to their risks and to intervene where appropriate. This interaction is intended to foster an active dialog between banks and supervisors such that when deficiencies are identified, prompt corrective action can be taken to either reduce risk or restore capital.

There are three main areas that might be particularly suited to treatment under Pillar 2: risks considered under Pillar 1 that are *not* fully captured by the Pillar 1 process (e.g. credit concentration risk), factors not accounted for by Pillar 1 (e.g. interest-rate risk in the banking book, business and strategic risk), and factors external to the bank (e.g. business cycle effects). Moreover, Pillar 2 also involves an assessment by regulators of compliance with the minimum standards and disclosure requirements of the more advanced methods in Pillar 1, such as the IRB framework for credit risk and the AMA for operational risk.

Four Key Principles of Supervisory Review: Basel II identifies four key principles of supervisory review, which are described in Figure 12.6 below.

In addition to these principles, Basel II cautions bank supervisors to carry out their obligations in a transparent and accountable manner. Moreover, it encourages enhanced cooperation between national supervisors, especially for the cross-border supervision of complex international banking organizations.

The Third Pillar: Market Discipline

Given the increasing complexity of banking activities, it is extremely difficult, if not impossible, for banking supervisors to monitor these activities in detail. Basel II therefore encourages monitoring of banks by professional investors and financial

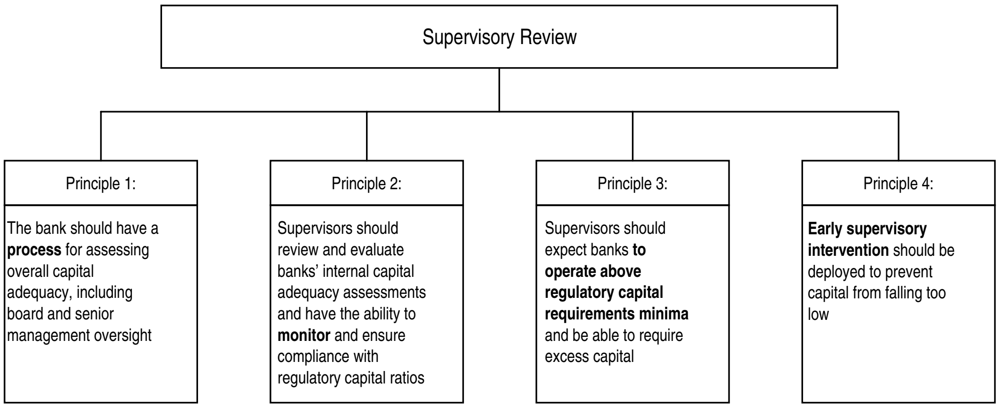


FIGURE 12.6 Key Principles of Supervisory Review

analysts as a complement to banking supervision. However, this is where Basel II provides the least detail and precision. Other than emphasizing the need for increased transparency, it says little about how to implement Pillar 3. This is somewhat surprising since a number of authors have provided very specific proposals on this issue.[41]

A Brief Evaluation of Basel II

Basel I was a major improvement over what preceded it. It was admittedly too crude for the purists and it introduced a wedge between the market assessment and regulatory assessments of asset risks. However, Basel II may have gone overboard in attempting to deal with all of the risks modern banks face. It has consequently become quite complex, and much of the intuitive, common-sense appeal of Basel I appears to have been lost. And despite this complexity, Basel II is silent about key issues. For example, it insists greatly on the need to "enable early supervisory intervention if capital does not provide a sufficient buffer against risk," and it devotes considerable effort to the design of risk weights, and yet is silent on the threshold and form of intervention.[42]

Other concerns are that the capital requirements under Basel II are likely to be *procyclical*.[43] The reason is that the accord seeks to link capital requirements more closely to risks, so that capital requirements will increase when risks increase. However, the risks are not observable and must be estimated. Standard methods to estimate risks tend to provide higher estimates during economic downturns; for example, during downturns we typically observe an increase in the number of credit rating downgrades and/or a widening of credit quality spreads. Of course, it is not obvious that risks actually increase during downturns. In fact, they may actually increase during upturns when the imbalances that may generate the next recession are building up and some lenders are lending as if the economic boom will be permanent.[44]

41. See Bliss (2001), Calomiris (1998), and Evanoff and Wall (2000).

42. See DeCamps, Rochet and Roger (2004), who make this point. An analysis of the effectiveness of the risk weights in the "Standardized Approach" appears in Resti and Sironi (2007).

43. See Ayuso, Pérez and Saurina (2004).

44. See Thakor (2005), for example.

So, what should an optimal design of capital requirements look like? Here we agree with many scholars who believe that capital requirements should focus less on complex schemes designed to deliver precision of measurements and more on acting as an intervention threshold for banking supervisors.[45] That is, capital requirements should focus on what regulators should do when banks do not comply with capital requirements.

The Debate Over Capital Requirements

There is an active debate about how high bank capital requirements should be. We provided part of the discussion of this issue in the previous chapters. Here we extend that discussion.

It is useful to begin by reiterating that regulatory capital requirements make sense only when there is a governmental safety net provided through deposit insurance, TBTF policies and other interventions. The argument in favor of such a safety net has both *ex ante* and *ex post* elements. The *ex ante* argument is that it minimizes the *likelihood* of bank runs and hence improves banking stability and the willingness of banks to make risky loans that are essential to economic growth. The *ex post* argument is that if a few banks fail, deposit insurance prevents the failures from spreading through the banking system. Of course, once such a safety net is implemented, the loss of market discipline and the moral hazard associated with excessive risk-taking by banks is inescapable.

Regulatory capital requirements are essentially a response to this moral hazard. They can be viewed as an attempt by regulators to "have their cake and eat it too." Capital requirements permit us to capture the ex ante and ex post benefits of the deposit insurance safety net without the pernicious effects of this safety net on banks' risk-taking incentives. This would seem to suggest that capital requirements should be set quite high.

However, here is where the discussion gets murky. There are many who argue that increasing capital requirements has many undesirable effects. These are that beyond a certain point a higher capital requirement may: (i) induce banks to take *more* risk; (ii) cause banks to rely less on deposits, thereby diminishing liquidity creation by banks; and (iii) impose adverse selection costs on banks as they attempt to raise additional equity to meet the higher capital requirements. We evaluate each in turn.

Regarding (i), whether capital requirements at the current levels are high enough to produce this perverse effect is unclear. We do not believe this to be so, but will return to this question shortly. Moreover, this argument applies only in a rather narrow set of circumstances, even theoretically. As for (ii), the argument has been made in many different ways by different authors.[46] One variant is that if one fixes the total size of the bank, then the more capital a bank has, the less deposits it has. That is, bank equity capital "squeezes" out deposits. This is tautological.[47] Consequently, there is less liquidity transformation as the bank makes loans that are funded with less deposits. Another variant points to the withdrawability of deposits, which creates market discipline for the bank and disciplines it on its asset investments (recall

45. See Dwatripont and Tirole (1994) and DeCamps, Rochet and Roger (2004).

46. See, for example, Diamond and Rajan (2001) and Gorton and Winton (2000).

47. One can drop the fixed-size assumption and still make a similar argument that is not tautological. In a general equilibrium setting, Gorton and Winton (2000) argue that higher bank capital requirements induce a substitution effect from bank deposits to bank equity.

our discussion in Chapter 10). Depositors, aware of this, will then be willing to provide funding to the bank. Equity, on the other hand, does not have this withdrawability feature, and hence lacks the disciplining effect of deposits. This means that bank's shareholders may withhold funding in instances in which deposit funding may have been viable. This leads to lower asset investments (loans) by banks and smaller bank balance sheets. Thus, increasing bank capital requirements could lead to lower lending and less liquidity creation by banks. In other words, higher bank capital requirements may reduce qualitative asset transformation. This implies that higher bank capital requirements also lower bank profitability as banks either shrink their balance sheets or invest less in risky loans and more in lower-margin marketable securities where the asset-investment-discipline issues are less a concern for the bank's shareholders.

As for (iii), the basic argument is that bank equity is more prone to adverse-selection costs induced by asymmetric information than deposits or subordinated debt.[48] As capital requirements increase and banks issue additional equity in response, they incur higher adverse-selection costs. The empirical prediction is that banks become less profitable as they keep more capital. However, we lack a well-articulated theory of why *bank* capital should be especially prone to this adverse-selection cost.

What does the empirical evidence say? Here are the findings.[49] First, more stringent capital regulation reduces nonperforming loans and contributes positively to bank development, although the link between capital regulation stringency and development weakens significantly when one controls for other features of bank supervision and regulation. Second, generous deposit insurance schemes are strongly and *negatively* associated with bank stability.[50] Somewhat ominously, strong official supervisory agencies, stringent capital standards, and regulations that encourage private-sector monitoring of banks do *not* counterbalance these negative effects of generous deposit insurance. This *contradicts* the earlier argument we presented that imposing higher capital requirements can allow regulators to capture the benefits of the deposit insurance safety net without the adverse effects of deposit insurance. Third, regulations that encourage and facilitate private monitoring of banks lead to greater bank development and reduce nonperforming loans.

A recent empirical study also examines the effect of bank capital on the value of the target bank in an acquisition.[51] It finds that both the purchase price and the goodwill of the target bank are increasing in its equity capital. Moreover, the pre-acquisition operating performance of the target bank is also better when it has more equity capital.

There has also been recent work on the relationship between bank capital and performance that focuses exclusively on large U.S. banks.[52] Constructing a scorecard based on return on average assets, return on average equity, Tier-1 capital ratio, leverage ratio, the nonperforming asset ratio and the reserve coverage (loan-loss reserves/total loans), the study identified the top 150 U.S. banks based on their performance during a four-quarter period spanning the last two quarters of 2004 and the first two quarters of 2005. Its principal finding is that the top-performing U.S. banks have very *high* capital levels. See Table 12.4, which identifies the top 25 banks.

We can summarize as follows. First, the continuing usefulness of federal deposit insurance is questionable. Second, higher levels of bank capital do not seem to

48. See Stein (2004), for example.
49. These findings are reported by Barth, Caprio and Levine (2004).
50. Also see Demirgüc-Kunt and Detragiache (2002).
51. See Mehran and Thakor (2007).
52. See Milligan (2005).

TABLE 12.4 Top 25 U.S. Banks in Operating Performance During 2004–2005

				Profitability				*Capital Adequacy*				*Asset Quality*				
Rank	**Company Name**	**State**	**Total Assets ($000)**	**Core ROA A (%)**	**ROA Rank**	**Core ROAE (%)**	**ROE Rank**	**Tier 1 Ratio (%)**	**Tier 1 Rank**	**Leverage Ratio(%)**	**Leverge Rank**	**NPAs/ Loans & OREO (%)**	**NPA Rank**	**Reserves/ Loans (%)**	**Reserve Bank**	**Final Score**
1	Bank of Hawaii Corp.	HI	10,059,690	1.85	9	25.79	3	10.25	80	7.42	95	0.18	17	1.65	12	114.0
2	S & T Bancorp	PA	3,095,177	1.96	7	17.22	37	10.29	78	9.68	22	0.29	36	1.40	33	128.5
3	Park National Corp.	OH	5,633,319	1.74	12	17.72	32	14.14	18	8.94	33	0.67	114	2.14	5	129.0
4	City National Corp.	CA	14,475,598	1.63	18	16.82	41	11.91	43	8.12	62	0.25	32	1.66	11	133.0
5	Glacier Bancorp	MT	3,531,935	1.52	32	18.07	30	12.06	38	8.86	35	0.34	53	1.53	23	136.5
6	Westcorp	CA	16,544,234	1.60	21	17.88	31	11.09	57	8.75	39	0.44	78	2.49	1	139.5
7	Commerce Bancshares	MO	14,118,193	1.53	29	15.48	55	11.77	45	9.46	26	0.20	22	1.52	25	143.0
8	CVB Fin'l Corp.	CA	4,811,854	1.46;	42	21.28	9	12.10	34	8.27	54	0.00	1	1.05	96	143.5
9	Mellon Fin'l Corp.	PA	36,935,000	2.23	1	19.70	16	10.85	62	8.06	63	0.35	58	1.15	73	145.0
10	Cullen/Frost Bankers	TX	9,950,973	1.59	23	18.46	24	12.84	23	9.51	25	0.74	121	1.38	35	149.0
11	Synovus Fin'l Corp.	GA	26,713,294	1.97	6	18.62	22	9.95	92	9.65	24	0.50	90	1.34	41	151.5
12	Westamerica Bancorp.	CA	5,191,093	2.09	4	25.12	4	9.04	113	5.96	144	0.29	36	2.23	3	156.0
13	Corus Bankshares	IL	6,504,221	2.05	5	20.36	12	12.08	35	14.19	4	0.82	130	0.94	111	157.0
14	UCBH Holdings	CA	7,037,174	1.53	29	20.05	14	11.98	40	9.10	30	0.45	82	1.13	80	159.0
15	First Midwest Bancorp	IL	7,073,141	1.52	32	19.80	15	10.31	77	8.18	59	0.34	53	1.33	44	163.5
16	Mercantile Bankshares Corp.	MD	16,092,994	1.72	13	12.85	86	11.49	48	10.18	17	0.24	31	1.38	35	164.5
17	U.S. Bancorp	MN	203,981,000	2.22	2	22.61	7	8.10	135	7.90	72	0.46	85	1.54	22	166.0
18	Hancock Holding Co.	MS	4,789,065	1.51	35	15.24	57	12.41	29	8.75	38	0.35	58	1.45	27	168.0
19	International Bancshares Corp.	TX	10,273,970	1.30	60	17.39	34	11.69	46	6.80	118	0.03	3	1.67	10	182.5
20	Cathay General Bancorp	CA	6,073,302	1.64	16	13.77	73	10.60	69	9.00	31	0.36	62	1.47	29	183.0
21	East West Bancorp	CA	6,701,584	1.50	36	18.13	28	9.61	100	8.83	37	0.04	4	0.99	104	186.5
22	UnionBanCal Corp.	CA	51,178,058	1.58	25	18.90	20	8.88	117	7.79	79	0.21	26	1.22	62	187.0
23	Comerica	MI	54,689,000	1.72	13	17.38	36	8.52	127	10.50	16	0.57	104	1.41	30	187.5
24	Washington Federal	WA	7,930,456	1.77	11	11.85	98	26.25	4	14.37	2	0.15	12	0.43	140	188.0
25	Hanmi Fin'l Corp.	CA	3,251,792	1.86	8	14.12	69	11.22	54	9.41	27	0.23	30	0.91	115	190.0

Source: Bank Director, pp. 17–18, 4th Quarter 2005.

adversely affect bank performance. In fact, both the operating performance and the value of the bank appear to be higher when the bank has more equity capital. However, if deposit insurance is phased out, the case for bank capital *requirements* is weakened. Third, expanded banking powers seem to have positive economic outcomes owing to increased competition.

Conclusion

U.S. banking history was shaped by American populism and the frontier mentality. The result was a fragmented financial services industry and a similarly fragmented public regulation of financial services. The issuance of bank notes and the need for cash-asset reserves conditioned the focus on liquidity in banking. The pyramiding of liquidity reserves led to systemic risk in the national banking system of the 19th century and to periodic financial panics. This led to the creation of the Federal Reserve System. Federal deposit insurance was added in 1933 in response to a virtual collapse of the banking system.

Regulation of interest rates and controlled entry into banking created monopoly rents for banks and mitigated the moral hazard arising from deposit insurance. This provided stability that lasted until the late 1970s when higher and more volatile interest rates induced massive disintermediation. Regulators responded by lifting interest-rate ceilings, relaxing investment restrictions, and reducing regulatory scrutiny. Thrifts that were legally locked into fixed-rate mortgages sustained huge losses owing to the consequent interest-rate risk. These losses impelled thrifts to undertake greater credit risk, resulting in further losses. The implosion of the thrift industry eventually led to a series of legislative and regulatory initiatives including FIRREA (1989) and FDICIA (1991). This was subsequently followed by new capital requirements under Basel II.

Regulatory reform has almost always been a reflexive reaction to some financial crisis. But regulation should be forward looking and future banking regulation should consider the key issues of optimal risk-taking and innovation by financial intermediaries.

Review Questions

1. What are the key milestones of bank regulation in the United States?
2. If deposit insurance is deemed necessary, what steps should be taken to reform the system?
3. What are the pros and cons of market value accounting?
4. What are the pros and cons of "expanded" banking powers?
5. Do regulators always maximize social welfare? Why or why not? Can anything be done about this?
6. Discuss the key elements of FDICIA and the NCFIRRE proposal. Provide a critique of each.
7. How would you reform our banking system?
8. Why did monetary policy initiatives during 1991–92 in the United States fail to move the economy out of its sluggish growth pattern?

References

Adams, James R., *The Big Fix: Inside the S&L Scandal*, New York: John Wiley & Sons, 1990.

Allen, Franklin, and Douglas Gale, "Measurement Distortion and Missing Contingencies in Optimal Contracts," *Economic Theory* 2, 1993, 1–26.

Ayuso, Juan, Daniel Perez, and Jesus Saurina, "Are Capital Buffers Pro-Cyclical? Evidence from Spanish Panel Data," *Journal of Financial Intermediation* 13–2, April 2004, 249–264.

Balderstone, Frederick E., *Thrifts in Crisis: Structural Transformation of the Savings and Loan Industry*, Cambridge, MA: Ballinger, 1985.

Barth, James R., Gerand Caprio, Jr., and Ross Levine, "Bank Regulation and Supervision: What Works Best?" *Journal of Financial Intermediation* 13–2, April 2004, 205–248.

Benston, George, *The Separation of Commercial and Investment Banking*: *The Glass-Steagall Act Revisited and Reconsidered*, New York: Oxford University Press, 1990.

Berger, Alan N., Kathleen K. King, and James M. O'Brien, "The Limitations of Market Value Accounting and a More Realistic Alternative," *Journal of Banking and Finance* 15, 1991.

Bhattacharya, Sudipto, and Anjan V. Thakor, "Contemporary Banking Theory," *Journal of Financial Intermediation* 3–1, October 1993, 2–50.

Boot, Arnoud W. A., "Why Hang on to Losers: Divestitures and Takeovers," *The Journal of Finance* 47–4, September 1992, 1401–1424.

Boot, Arnoud W. A., and Anjan V. Thakor, "Banking Scope and Financial Innovation," *Review of Financial Studies* 10–4, Winter 1997, 1099–1131, reprinted as a "CEPR Classic" in a book published by *Oxford University Press* (eds: Bruno Biais and Marco Pagasio).

———,"Ambiguity and Moral Hazard," Indiana University working paper, August 1991.

———, "Self-Interested Bank Regulation," *American Economic Review* 83–2, May 1993, 206–212.

Boot, Arnoud W. A., and Stuart I. Greenbaum, "Bank Regulation, Reputation and Rents: Theory and Policy Implications," in *Capital Markets and Financial Intermediation*, Colin Mayer and Xavier Vives, editors, Cambridge: Cambridge University Press, 1993, 262–285.

Boot, Arnoud W. A., Stuart I. Greenbaum, and Anjan V. Thakor, "Reputation and Discretion in Financial Contracting," *American Economic Review* 83–5, December 1993, 1165–1183.

Booth, James R., "FDIC Improvement Act and Corporate Governance of Commercial Banks," *Economic Review* 1, *Federal Reserve Bank of San Francisco*, 1993, 14–22.

Bray, Hammond, *Banks and Politics in America From the Revolution to the Civil Wars*, Princeton University Press, Princeton, N.J., 1957.

Campbell, Tim S., Yuk-Shee Chan, and Anthony Marino, "An Incentive-Based Theory of Bank Regulation," *Journal of Financial Intermediation* 2, 1992, 255–276.

Carnell, Richard S., "Implementing the FDIC Improvement Act of 1991," paper presented at a *Conference on Rebuilding Public Confidence Through Financial Reform*, Ohio State University, Columbus, Ohio, June 25, 1992.

Chan, Yuk-Shee, Stuart I. Greenbaum, and Anjan V. Thakor, "Is Fairly Priced Deposit Insurance Possible?" *The Journal of Finance*, 47–1, March 1992, 227–246.
Corrigan, Gerald F., "Reforming the U.S. Financial System: An International Perspective," *Quarterly Review of the Federal Reserve Bank of New York* 15, Spring 1990, 1–14.
DeCamps, Jean-Paul, Jean-Charles Rochet, and Benoit Roger, "The Three Pillars of Basel II; Optimizing the Mix," *Journal of Financial Intermediation* 13–2, April 2004, 132–155.
Demirgüc-Kunt, A., and Enrica Detragiache, "Does Deposit Insurance Increase Banking System Stability? An Empirical Investigation," *Journal of Monetary Economics*, 2005.
Dewatripont, Mathias, and Jean Tirole, *The Prudential Regulation of Banks*, MIT Press, 1994.
Diamond, Douglas, and Raghuram Rajan, "Liquidity Risk, Liquidity Creation and Financial Fragility: A Theory of Banking," *Journal of Political Economy*, April 2001.
Donaldson, R. Glen, "Costly Liquidation, Interbank Trade, Bank Runs and Panics," *Journal of Financial Intermediation* 2–1, March 1992, 59–82.
Eisenbeis, Robert A., and Paul M. Horvitz, "The Role of Forbearance and Its Costs in Handling Troubled and Failed Depository Institutions," paper presented at the American Finance Association meetings, Anaheim, CA, 1993.
Ely, Bert, "Surprise! Congress Has Just Enacted What Amounts to a Core Banking System," *American Banker*, September 27, 1993.
Fischer, Gerald C., *The Modern Bank Holding Company: Development, Regulation, and Performance*, Temple University, Philadelphia, PA, 1986.
Flannery, Mark J., "Pricing Deposit Insurance When the Insurer Measures Bank Risk with Error," *Journal of Banking and Finance* 15, 1991, 975–998.
Gorton, Gary, and Andrew Winton, "Liquidity Provision, Bank Capital and the Macroeconomy," mimeo, University of Minnesota, 2000.
Greenspan, Alan, "FDICIA and the Future of Banking Law and Regulation," *Proceedings of the 29th Annual Conference on Bank Structure and Competition, Federal Reserve of Chicago*, titled: "*FDICIA: An Appraisal*," May 1993, 3–9.
Huertas, T., "Can Banking and Commerce Mix?" *Cato Journal* 7, Winter 1988, 743–762.
Kane, Edward J., "Changing Incentives Facing Financial Services Regulators," *Journal of Financial Services Research*, 1989a, 265–274.
———, "The Unending Deposit Insurance Mess," *Science* 246, October 1989b, 451–456.
———, "Market Forces and Financial Misregulation," talk delivered at the *Third Annual Australasian Finance and Banking Conference*, November 1990b.
———, "Principal-Agent Problems in S&L Salvage," *The Journal of Finance* 45–3, July 1990a, 755–764.
Kane, Edward J., and M. Yu, "How Much Did Capital Forbearance Add to the Tab for the FSLIC Mess?" working paper, Boston College, 1993.
Kareken, John H., "Deposit Insurance Reform; Or, Deregulation Is the Cart, Not the Horse," *Quarterly Review, Federal Reserve Bank of Minneapolis*, Winter 1990, 3–11.
Kroszner, Randall, and Raghuran Rajan, "Is the Glass-Steagall Act Justified?" *American Economic Review*, forthcoming.

Loewy, Michael B., "The Macroeconomic Effects of Bank Runs: An Equilibrium Analysis," *Journal of Financial Intermediation* 1, 1991, 242–256.

Mehran, Hamid, and Anjan V. Thakor, "Does Capital Hurt the Value of the Bank?" Working Paper, Federal Reserve Bank of New York, January 2007.

Mester, Loretta J., "Banking and Commerce: A Dangerous Liaison?" *Business Review, Federal Reserve Bank of Philadelphia*, May/June 1992, 17–29.

Milligan, Jack, "What Makes a Good Bank Good?" *Bank Director*, 4th quarter 2005, 16–24.

Myers, Stewart C., and Nicholas S. Majluf, "Corporate Financing and Investment Decisions When Firms Have Information That Investors Do Not Have," *Journal of Financial Economics* 13, 1984, 187–221.

National Commission on Financial Institution Reform, Recovery and Enforcement; "Origins and Causes of the S&L Debate: A Blueprint for Reform," *A Report to the Congress of the United States*, Washington, D. C., July 1993.

O'Hara, Maureen, "Real Bills Revisited: Market Value Accounting and Loan Maturity, " *Journal of Financial Intermediation* 3–1, October 1993, 51–76.

Resti, Andrea, and Andrea Sironi, "The Risk-Weights in the New Basel Capital Accord: Lessons from Bond Spreads Based on a Simple Structural Model," *Journal of Financial Intermediation* 16, January 2007, 64–90.

Saunders, Anthony, "The Separation of Banking and Commerce," working paper, New York University Salomon Center, May 1991.

———, "Securities Activities of Commercial Banks: The Problem of Conflict of Interests, " *Business Review, Federal Reserve Bank of Philadelphia*, July/August 1985, 17–27.

Shull, B., "The Separation of Banking and Commerce: Origin, Development, and Implications for Antitrust, " *The Antitrust Bulletin* 28, Spring 1983, 255–279.

Simons, Henry C., *A Positive Program for Laissez Faire*, Public Policy Pamphlet No. 15, University of Chicago, 1934.

Smith, Rixey, and Norman Beasley, *Carter Glass: A Biography* (1939), New York: Da Capo Press, 1972 (reprint).

Stein, Jeremy, "Comment on Jean-Charles Rochet: Rebalancing the Three Pillars of Basel II" in *Federal Reserve Bank of New York Economic Policy Review* 10, September 2004, 27–29.

Thakor, Anjan V., "Do Loan Commitments Cause Overlending?" *Journal of Money Credit and Banking* 37–6, December 2005, 1067–1100.

Wall, Larry D., "Too-Big-to-Fail After FDICIA," *Economic Review, Federal Reserve Bank of Atlanta*, January/February 1993, 1–14.

White, Lawrence, "Market Value Accounting: An Important Part of the Reform of the Deposit Insurance System," in *Capital Issues in Banking*, Association of Reserve City Bankers and the Banking Research Center, Northwestern University, 1988.

PART ♦ VII

Overall Management of the Bank

This page intentionally left blank

CHAPTER ♦ 13

Management of Risks and Opportunities in Banking

"The only thing that saves you is capital. Good management just goes so far. Technology just goes so far. The regulators just go so far. It's capital."

Donald Powell, FDIC Chairman, 2005

Glossary of Terms

Duration-Adjusted Gap: A measure that attempts to capture the interest-rate risk inherent in assets and liabilities of different durations (see Chapter 4).

CMOs: Collateralized Mortgage Obligations (see Chapter 8).

HMDA: Home Mortgage Disclosure Act (see Chapter 2).

BSA: Bank Secrecy Act (see Chapter 2).

RESPA: Real Estate Settlement Procedures Act (see Chapter 2).

ATM: Automated Teller Machine.

TBTF: Too Big to Fail (see Chapters 9, 10, and 11).

LDCs: Loans to Developing Countries (see Chapter 12).

MSBs: Mutual Savings Banks (see Chapter 2).

L/Cs: Letters of Credit (see Chapter 7).

ROA: Return On Assets.

BIS: Bank for International Settlements.

Introduction

Purpose of Chapter: The purpose of this chapter is to draw together earlier, more detailed discussions of risks and opportunities facing financial institutions. We will focus upon banks, though many of the ideas can be applied to many other financial institutions as well. This chapter could have been the last in the book. We place it here because the chapters that follow may be included somewhat selectively in a one-term course.

Failures of Risk Management: Banks are in the business of managing risk and they seek ways to measure risk as the first step to controlling and pricing it. Some recent experiences of commercial and investment banks discussed below attest to the importance of risk management.[1]

- In February 1990, Drexel Burnham Lambert collapsed.
- In January 1991, the Bank of New England announced a possible loss of $0.5 billion in the fourth quarter of 1990. Depositors ran the bank, withdrawing $1 billion. The bank failed, costing the FDIC $2.3 billion.
- In December 1992, Barclays Bank announced a $375 million loan-loss provision against a $656 million loan to a British property developer (Imry). For 1992 as a whole, the bank set aside nearly $2.6 billion against bad loans.
- It is believed that J.P. Morgan lost nearly $200 million in 1992 trading mortgage-backed securities. Earlier, the investment banking house of Merrill Lynch lost $375 million in the same market.
- In January 1993, the German BFG Bank (a subsidiary of Credit Lyonnais) announced a $700 million loss for 1992. Much of the loss came from provisions against loans to developing countries.
- The Japanese government has assisted numerous banks that suffered massive losses on real estate assets and equity investments.
- In September 1992, the Australian bank, Westpac, reported a loss of $1.1 billion after writing off bad debts of $1.90 billion.
- In August 1998, Long-Term Capital Management, a hedge fund, collapsed with a reported loss of $1.85 billion in capital (more below).[2]

In addition to the above failures, we have also witnessed even more spectacular failures in recent years that can be attributed to a mix of risk-management failures and poor judgment, and in some cases outright fraud. These failures have given new meaning to the term "enterprise risk," which we alluded to in the previous chapter.

Each of these traumas is traceable to a failure of risk management. Many of these organizations simply failed to recognize the credit, liquidity, or interest rate risk that they faced. In other cases, they overextended themselves in trading securities.

The importance of risk management for financial institutions has been elevated by recent technological developments, the emergence of new financial instruments, deregulation, and heightened capital market volatility. For example, in the 1980s, banks began to enter discount brokerage, and in 1985, the Fed ruled that Bankers Trust could sell commercial paper. The overall management of risks has become

1. See "International Banking: A Comedy of Errors," *The Economist* (1993).
2. See Wikipedia.

The LTCM Debacle

One of the most spectacular of the risk-management failures was that of Long Term Capital Management (LTCM), which was a hedge fund founded in 1994 by the former vice chairman and head of bond trading at Salomon Brothers, Lee Meriwether. On its board of directors were Myron Scholes and Robert Merton, two finance professors who had earlier won the Nobel Prize in Economics for their work on option pricing models. LTCM began trading February 24, 1994, with a little over $1 billion in investor capital.

The company had developed complex mathematical models to take advantage of fixed-income arbitrage opportunities (termed *convergence trades*) usually with U.S., Japanese, and European government bonds. The basic idea was that over time the value of long-maturity bonds issued a short time apart would tend to converge. However, the rate at which the prices of these bonds converged would be different, and the more heavily traded bonds (such as U.S. Treasury Bonds) were expected to converge to the long-term price more rapidly than less heavily traded and less liquid bonds.

Thus, by a series of financial transactions (essentially amounting to buying the cheaper, less liquid bond and short selling the more expensive, but more liquid bond) it would be possible to make a profit as the difference in the values of the bonds narrowed.

As LTCM's capital base grew, the need for additional returns on that expanded capital caused it to undertake other trading strategies. Although these trading strategies were not dependent on market movements—they were not dependent on overall interest rates or stock prices going up (or down)—they were not "convergence trades" as such. By 1998, LTCM had extremely large positions in areas such as merger arbitrage and S&P 500 options.

Because these differences in values were relatively small – especially for the convergence trades – the fund needed to take highly leveraged positions in order to make a significant profit. At the beginning of 1998, the firm had equity of $4.72 billion and had borrowed over $124.5 billion with assets of around $129 billion. It had off-balance sheet derivative positions of $1.25 trillion, most of which were in interest-rate derivatives such as interest-rate swaps. The fund also invested in other derivatives such as equity options.

The downfall of the fund started in May and June 1998 when net returns fell to −6.42 percent and −0.14 percent, which sapped LTCM's capital by $461 million. The situation was further exacerbated by the exit of Salomon Brothers from the arbitrage business in July 1998.

Things finally came to a head in August and September 1998 when the Russian government defaulted on their government bonds (GKOs). Panicked investors sold Japanese and European bonds to buy U.S. Treasury bonds. The profits that were supposed to occur as the value of these bonds converged became huge losses as the values of the bonds diverged instead of converging. By the end of August, the fund had lost $1.85 billion in capital.

LTCM, which was providing annual returns of almost 40 percent up to this point, experienced a "flight to liquidity." This prompted a bailout of $3.625 billion by the banks, organized in September 1998 by the Federal Reserve Bank of New York, arguably to avoid a wider collapse in the financial markets. The fear was that there would be a contagion effect as the company liquidated its securities to cover its debt, leading to a drop in prices that would force other companies to liquidate their own debt, further depressing prices.

The total losses were $4.6 billion.

How did the nature of the risks taken by LTCM lead to its eventual demise, and were its founders wrong in their assumptions? To address this question, we begin by noting that the profits from LTCM's trading strategies were generally not correlated with each other and thus normally LTCM's highly leveraged portfolio benefited from diversification. However, the general flight to liquidity in the late summer of 1998 led to a marketwide repricing of all risk and these positions then did all move in the *same* direction. As the correlation of LTCM's positions increased, the diversified aspect of LTCM's portfolio vanished and large losses were suffered by its shareholders.

In the end, LTCM's problem turned out not to be its basic idea, which was correct; the values of government bonds of different liquidities did eventually converge. But this happened only after the company was wiped out. Nonetheless, the incident confirms an insight often attributed to the economist John Maynard Keynes, that although markets do tend toward rational positions in the *long* run, "the market can stay irrational longer than you can stay solvent."

The fall of LTCM is an important example of the principle that arbitrage is *not* riskless (see Chapter 1).

Source: Wikipedia.

substantially more complex, with banks now being allowed to do more and more subsequent to the passage of the Gramm-Leach-Bliley Act in 1999.

Each new opportunity brings with it risks that place greater demands on bank management. But the causality runs the other way too. To cope with expanded risks, banks must develop new instruments and enter new markets. This creates new profit opportunities. The growth of securitization (Chapter 9) illustrates this point.

Organization of Chapter: The rest of the chapter is organized as follows. The next section reviews the various risks and opportunities faced by banks. Much of this discussion synthesizes selected risk-related material from previous chapters. In the section that follows, we discuss the normal day-to-day management of risks and opportunities, that is, issues that bank management must address on an ongoing basis. Included in this is a discussion of interbank payment systems. After that we take up crisis management, that is, the management of risks of potentially catastrophic magnitude. The subsequent section is devoted to strategic planning, that is, the management of options that shape the bank's long-term profitability. This is followed by some case studies on the management of risks and opportunities.

Opportunities and Risks in Banking

In this section, we provide an overview of the opportunities and risks faced by a bank. In previous chapters, we have detailed the different risks banks confront. We now draw selectively upon that material.

The Evolution of Opportunity and Inherent Risks for a Depository Intermediary

You will recall from Chapters 2, 3, and 4 that in providing transactions and other services to depositors and in screening and monitoring loans, the bank unavoidably processes credit and liquidity risks. Even our primordial goldsmith, after evolving into a rudimentary bank, faces liquidity risk stemming from unanticipated withdrawals of gold deposits. Moreover, when he lends warehouse receipts to merchants, he accepts credit risk (Chapters 5 and 6). And since the concerns of depositors about the goldsmith affect withdrawal patterns, leading periodically to bank runs and panics (Chapter 10), a correlation arises naturally between liquidity and credit. We emphasized the importance of informational frictions. If all information about the bank's loans were costlessly available, the bank could liquidate assets with little sacrifice in value in order to meet withdrawal demands, and liquidity risk would be trivialized.

Interest rate risk also can be understood in the context of the goldsmith story. Because deposit withdrawal patterns cannot be matched precisely with loan maturities, the goldsmith who lends warehouse receipts is exposed to a duration-mismatched balance sheet that will create unpredictable gains and losses from interest rate fluctuations.

Thus we see that credit, liquidity, and interest rate risks arise quite naturally from the evolution of the goldsmith into a fractional reserve bank. Note that this evolution can be traced to the goldsmith's recognition of the *profit opportunities* inherent in the social acceptance of warehouse receipts as a medium of exchange. The correlation between opportunity and risk is therefore evident as the goldsmith steps across the line from being a mundane warehouse operator to being a fractional reserve banker. This takes place the moment the goldsmith prints receipts in excess of his gold holdings.

As the primitive banker became more sophisticated, he began to realize other profit opportunities. There were times when the banker had to turn down loan requests from merchants because it would have required printing more warehouse receipts than he wished to. So the goldsmith-turned-banker asked: "How can I profit from this?" The answer was: by *selling* promises of future loans to merchants in advance of their actual borrowing needs. Then the banker could collect fees even before he had to lend, and the merchants would be happy because they were assured of future access to credit. Thus was born the bank loan commitment (Chapter 8)! Of course, the banker soon realized that this practice carried risks as well; there were times when the banker had insufficient warehouse receipts to make good on his promises or he regretted *ex post* having promised to lend at terms that seemed too generous when the merchant actually came for the loan. These risks also had to be managed.

The major risks faced by banks changed very little until the latter half of the 20th century. Most banks had adequate capital and operated in a low-volatility environment characterized by fixed exchange rates, stable interest rates, capital controls, and oligopolistic credit markets. Consequently, risk management techniques did not change very much.

Things began to change rapidly, however, when market volatility increased and regulatory protection for financial institutions began to diminish. In the mid-1960s, the earnings of thrifts—S&Ls and MSBs—began fluctuating widely as their cost of funds changed much more rapidly than their earnings on long-term fixed-rate mortgages. We have seen in previous chapters how this interest-rate risk, in combination with a host of credit risk problems, led to large-scale thrift failures.

During this time, commercial banks managed interest-rate risk more successfully, but encountered other problems. Before 1970, banks were either intensively regulated or comfortably cartelized in most industrialized countries. In 1971–72, however, the Bretton Woods agreement was dismantled and currencies were allowed to float freely; this elevated foreign-exchange risk. Then came the oil-price shocks, high inflation, and wild movements in interest rates. Risk in banking grew significantly both in variety and magnitude.[3] Moreover, as the risks exposure of banks' customers increased, they increased their demand for risk-management services.

Banks responded with an array of new products that facilitated risk management: variable-rate loan commitments, standby letters of credit, swaps, options futures, and more. These product developments were also spurred by two important phenomena. One was the explosive growth in exchange-traded derivatives (see the discussion in Chapter 8), and there has been even more growth in those traded in the over-the-counter (OTC) markets where maturities tend to be longer than those of derivatives traded on organized exchanges. The other was the acceleration in the power of computers, which gave banks new ways to analyze contracts, customers, and markets and a significantly expanded capability to handle payments. Having developed a variety of products to satisfy this demand, bankers learned that they could make loans without having to fund them, thus giving rise to securitization and loan sales (Chapter 9). We refer to the risks of these activities as *off-balance sheet*, but they nevertheless involve credit risk, liquidity risk, and interest-rate risk.

Where the management of risks fits into a bank's overall financial management is shown in Figure 13.1.

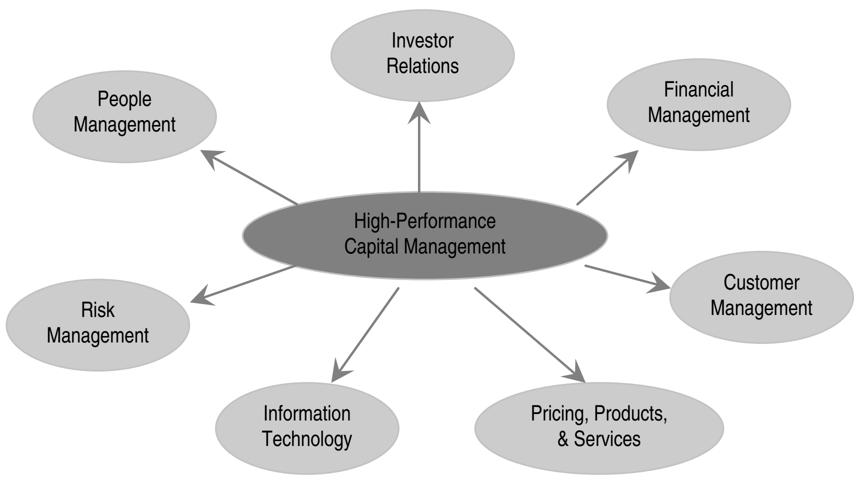


FIGURE 13.1 Risk Management As a Part of Financial Management
Source: *Bank Director* 3rd Quarter 2005.

3. See Pierce and Chase (1988).

Other Risks

The risks faced by even a very simple bank can result in failures. It may be socially desirable to limit such failures. One reason is that the resulting dislocations may adversely affect economic activity. Bank failures destroy circulating media of exchange (money) and rupture bank-customer relations, interrupting access to credit.[4]

In the interest of economic stabilization and growth, the government may choose to intervene by lending against the collateral of the bank's illiquid loan assets. Or, the government could insure deposits (Chapter 10). But the lender-of-last-resort (LLR) facility and deposit insurance create moral hazard problems because the bank finds it advantageous to accept greater amounts of the three basic risks. The rational banker will reduce its holding of cash-asset reserves with the introduction of a lender-of-last-resort. Why screen credit risks and monitor borrowers as carefully when depositors are oblivious to the bank's liquidity and credit risks?

Cognizant of these distortions, the government will find it compelling to regulate banks (Chapter 11). Thus, regulation arises quite naturally as the goldsmith matures into a banker. But regulation creates risks of its own for banks. These are of two types. First, the government's perception of how best to regulate and deal with the moral hazard will be based on experience with the entire banking industry, including most notably the industry's most aberrant members. Since an individual bank cannot control or even predict the behavior of the industry, *changes* in regulation become *unpredictable* for individual banks, even when the government's reaction to a change in industry behavior is perfectly predictable. This unpredictability is magnified if the regulator randomizes its own strategy or employs ambiguity as an instrument of public policy.

A second regulatory risk stems from subsidies imbedded in the safety net. For example, the availability of the LLR facility at any finite interest rate involves a subsidy if the bank finds it optimal to expand its lending after the introduction of the LLR. This subsidy obligates the banks, and the government may wish to extract concessions in order to advance socioeconomic and political agendas. The Community Reinvestment Act codifies this viewpoint. Banks have special obligations to provide credit facilities to their local market areas because of the special privileges and protections accorded via legislation and public regulation. Senator Howard Metzenbaum made the point in a comment made before the House Banking Committee:[5]

> "Enough is enough. Let's tell the financial institutions that this gimmie-gimmie game is over. It's time the banks and savings and loans gave something back. It's time they addressed the banking needs of elderly and low-income Americans."

The combination of these two sources of risk may be called *regulatory risk* for a bank. In the terminology of asset pricing theory, this is a *nondiversifiable* risk (Chapter 1) and can be expected to increase the bank's cost of capital.

4. See Besanko and Thakor (1993), who argue that increasing interbank competition may hurt borrowers because it makes bank failures—and hence the rupturing of bank-customer relationships—more likely. Boot and Thakor (2000) show that greater interbank competition causes the *nature* of relationship lending to change. Song and Thakor (2006) argue that greater competition can change the *mix* of financing—core deposits versus purchased money—used by the bank.

5. See *Consumer Access to Basic Financial Services* (1989).

Finally, banks face a variety of universal business risks such as the risk of fraud, professional liabilities, the risk of physical hazards, and other calamities. All of these risks require management, but few are unique or even central to banking.

A Taxonomy of Risk and Opportunity Management

The management of risks and opportunities can be classified as: day-to-day management, crisis management, and strategic planning. By *day-to-day management*, we mean the continuous monitoring and management of the risks discussed thus far and the routine management of opportunities related to existing markets and relationships with existing customers. Crisis *management* deals with infrequent adverse events such as hostile takeovers, large-scale fraud, and similar misfortunes. And *strategic planning* involves the development of plans for the exploitation of new markets and opportunities and for the control of the attendant risks. Figure 13.2 provides a graphic of this taxonomy.

Day-to-Day Management

In this section, we discuss the management of: credit risk, liquidity risk, interest-rate risk, and off-balance sheet risk. We then address regulatory and miscellaneous other risks. We end the section with a discussion of the management of routine opportunities.

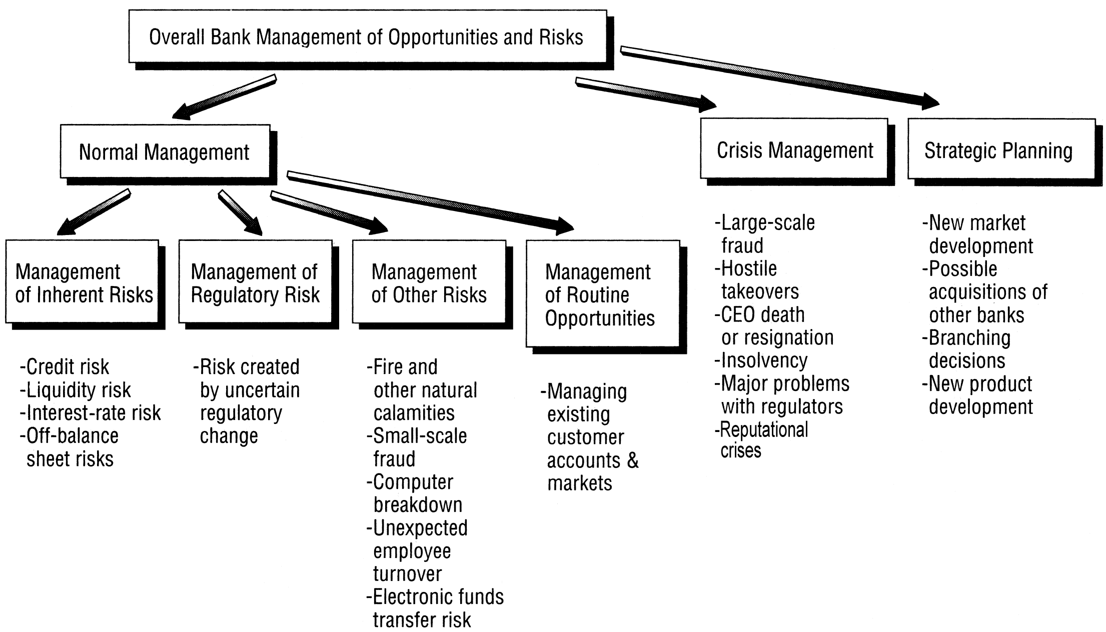


FIGURE 13.2 Management of Risks and Opportunities in a Bank

Management of Inherent Risks

Credit Risk: Management of credit risk is important because U.S. commercial banks obtain the bulk of their operating income from interest on loans. Loan losses represent changes in credit risk. Figure 13.3 shows how loan losses have changes over time for U.S. banks. After reaching a peak in 1991–92, there was a decline through 1994, after which there was again a surge. This reflects a shift by banks to higher credit risk.

Active and consistent application of sound credit policies can provide the sustainable competitive advantage that is so elusive in a "commodity" business, that is, one in which services are easily copied by competitors. Loans typically represent 60 percent or more of the bank's assets (recall Chapter 5) and capital is usually less than 15 percent. Consistently reducing credit losses by 1/2 percent translates into a much larger return on equity. From another vantage point, note that well-performing banks earn about 1 percent after taxes on their assets. A 1/2 percent reduction in losses on 70 percent of assets can increase ROA by 30 percent. This is the fundamental arithmetic of credit!

Managing credit risk involves four kinds of activities: underwriting or loan origination, funding and servicing the loan, risk processing (which includes monitoring and diversification), and sustaining the credit culture. We have discussed much of this in Chapters 5, 6 and 7, so we will be brief here.

(a) *Underwriting* or *Loan Origination*: Underwriting or loan origination embraces all those activities that precede the consummation of the loan, or alternatively, the rejection of the loan application. This is the "blocking and tackling" of credit. It includes credit analysis as well as design of the loan contract, which includes covenants, collateral, terms, and price. The bank's credit-granting decision is guided by its credit analysis as well as by its *written policy statement*, which is approved by the bank's board of directors and reviewed by bank examiners.

A good loan policy statement clearly states the bank's goals and guidelines for credit-risk management. It should also provide the following:

- The aggregate amount of loans the bank should make.
- The geographical areas from which loan business should be solicited.
- Limits on maximum loan sizes to various types of borrowers.

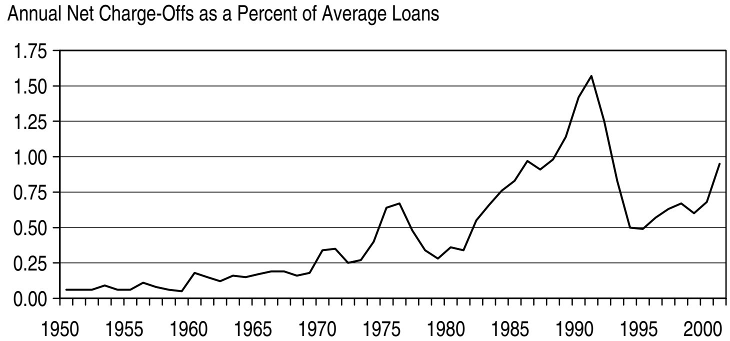


FIGURE 13.3 Rising Loan Losses Reflect a Gradual Shift to Higher Credit Risk in Banking (FDC-Insured Commercial Banks)
Source: FDIC Historical Statistics on Banking.

- The bank's diversification objectives.
- The maximum acceptable loan maturities.
- Criteria for collateral requirements.
- Standards of credit analysis and legal documentation.
- The appropriate relationship between credit risk and the price of the loan.
- The lines of authority and responsibility in the management of credit risk, that is, the relationship between the size of the loan and the level at which it can be approved.

(b) *Funding and Servicing the Loan: The Role of Documentation*: For the above guidelines to translate into effective credit-risk management, proper documentation is necessary; it is an important part of underwriting as well as *funding* and *servicing* the loan. When the Continental Illinois Bank failed in 1984, it was quickly discovered that hundreds of loan files, particularly those originated by Penn Square Bank in Oklahoma, were vacuous.[6] This was followed by the departure of Continental's internal auditor (who, without delay, found similar employment with another sizeable commercial bank in Chicago) and a lawsuit (eventually lost) against the external auditor by the FDIC. The documentation issue would again seem to be a routine matter, but during periods of rapid expansion, ordinary disciplines are often suspended, especially when the credit culture is not well established.

Documentation protects the bank's legal claim against the borrower for loan repayments. It also protects the bank's lien on the collateral. Finally, documentation assists the bank in monitoring the loan as to compliance with covenants.

(*c*) *Risk Processing*: After the loan is made, *risk processing* becomes important. We shall focus here on the *monitoring* and *diversification* aspects of risk processing. Monitoring subsumes all facets of the bank-customer relationship during the time interval between extension and complete repayment of the loan. In this period of indebtedness, the bank is a partner in the business. Monitoring involves following the borrower's fortunes (staying in touch), policing compliance with loan covenants, and managing the credit when covenants are violated.

A disciplined approach to monitoring is the product of an effective credit culture. Monitoring is a tool of credit management that limits the bank's exposure to credit losses. It does not necessarily mean "pulling the plug" and forcing default at the first opportunity. Rather it means devoting sufficient resources to identifying the first opportunity so that conscious choices can be made to protect the bank's interests. Of course, too much of a good thing is also possible when it comes to devoting resources. Monitoring is costly, so that there is an optimization that must be addressed. But redundant monitoring seems to have been far less a problem than the reverse in recent bank experience.

Restructuring and workouts are in the domain of monitoring. These are the monitoring activities for loans that have gone away. Responsibility for monitoring is typically shared between the bank's credit professionals and the calling officers (marketing specialists responsible for maintaining the customer relationship). This sharing of responsibility is one of the more delicate credit management issues.

6. In the case of Continental, there were other reasons as well for the failure of the bank. Even the numerous problems that were documented were ignored by the bank.

The calling officers prefer forbearance in order to preserve the customer relationship, whereas the credit professionals tend to be more aggressive in enforcing covenants since their responsibility is to minimize credit losses. Here again, the credit culture of the bank, defined and administered at the very highest management levels, must assure consistent application of policy.

Effective management of credit risk should also focus on *diversification* in the loan portfolio. The importance of diversification was brought home quite forcefully in the 1980s when numerous banks failed due to heavy loan concentrations in farming, energy, and real estate.[7] Banks sometimes approach diversification as a *constraint* rather than as an *opportunity*. That is, they establish techniques to identify "excessive concentrations" in loans with similar attributes, and they then try to reduce these concentrations. However, a good diversification program should be more aggressive—it should specify that loans should be sought and in what quantities. Moreover, it should prescribe *pricing differentials* as a way to encourage loan portfolio diversification. A loan that reduces portfolio risk should be priced lower than its own risk-adjusted interest rate would suggest, whereas one that raises portfolio risk should be priced higher. In other words, banks should take portfolio approach to credit risk, and use this approach to determine not only the prices of loans but also how much capital to allocate to them.

Of course, the benefits of diversification should be balanced against the benefits of *specialization* in determining the optimal degree of diversification. Specialization leads to more efficient credit analysis and the development of better monitoring techniques because of the cross-sectional and intertemporal information reusability.[8] But, as we discussed in Chapter 9, securitization can enable the bank to "have its cake and eat it too"; the bank can specialize in loan origination and monitoring and then diversify by securitizing some of its own loans and buying the securitized loans of others.

Apart from the trade-off between information reusability and diversification, processing credit risk involves a second more basic problem. Unlike the investor in traded bonds or equity who can array a wide variety of investment opportunities at each moment in time and reconfigure the portfolio at will, the banker's opportunities arrive sequentially and the ability to trade loans is limited (recall Chapter 6). Thus, each lending decision is made not only in the light of past decisions, but also with regard to uncertain future opportunities. This makes the lending policy a dynamic programming problem, and therefore inherently more complicated than the linear programming representation often used to characterize investment decisions where claims are tradable. Given the bank's finite capacity, defined by its human and financial capital, and the uncertain arrival of future lending opportunities, each lending decision must respond to the question of whether the opportunity is "beautiful enough." This framing of the lending decision has typically biased the banker to view each opportunity as a "tub on its own bottom," rather than an element in an ideally diverse portfolio. Thus, assessing risk concentrations by

7. Pierce and Chase (1988) quote a prominent banker as once saying: "Diversification is the only free lunch in banking." Also recall our discussions in Chapter 5.

8. Empirical evidence on the benefits of specialization has been provided by Eisenbeis and Kwast (1991). They found that over the 1978–88 period, the *average* earnings performance of banks specializing in real estate was on par with regular commercial banks, and those that had specialized in real estate for a longer period had higher returns with less risk than significantly-more-diversified commercial banks.

geography, industry, and product type is not enough. The banker must incorporate the dynamic change factor.

(*d*) *Credit Culture*: Being able to integrate underwriting, funding and servicing, and the risk processing activities of monitoring and diversification into an appropriate credit management system requires the appropriate *credit culture*, a term used to describe the bank's general approach to managing its credit risk. At the level of organizational design, the credit culture is defined by whether the board of directors has a credit committee (not all banks have one), and how that committee is empowered by the board. Then there is the question of the status and reporting responsibilities of the bank's chief credit officer (CCO). Does the CCO report directly to the board and its credit committee, much as the internal auditor reports to the board's audit committee, or does the CCO report directly to the CEO? Is the CCO in the top tier of management, or relegated to the second or third tier of management? Answers to these questions define the priority that the board of directors assigns to credit-risk management. Once these issues of priority are established, the formulation, articulation, and administration of lending policies come to the fore. A written loan policy, discussed earlier, is the next step.

Perhaps most important in the development of the credit culture is its most intangible and elusive element: ensuring the consistent and disciplined application of the bank's credit standards. It requires great discipline to resist the temptation to be "part of the herd" and to be myopic.[9] The fundamental problem of credit culture is ensuring that at moments of greatest temptation the whole credit apparatus is not pushed aside in a faddish feeding frenzy. Examples of herding behavior litter banking archives: REITS, the LDCs, LBOs, and commercial real estate are among the more recent forays. Even the exceptions are instructive. For example, Continental Illinois Bank publicly avoided LDC lending on the sensible argument that the spreads offered did not adequately compensate lenders for the credit risk.[10] Rather, continental chose to concentrate, and found its Waterloo in the oil patch of Oklahoma and Texas.

The relearning of credit principles has been slow and checkered and remains far from complete.[11] The *American Banker*, the daily trade paper of U.S. banking, titillates readers regularly with stories of major banks discovering the need for a strong credit culture. How could this be possible, given the centrality of credit-risk management? Apparently, many banks even now continue to manage as if the lower-risk strategies of the past remain viable. Of course, some banks have been quite successful in credit-risk management. In the box below we provide some examples.

9. Rajan (1994) explains some of the temptations bankers face to "follow the herd." See also Thakor (2005).

10. To this, Walter Wriston, then CEO of Citicorp, countered that countries do not go bankrupt. True perhaps, but it does not follow that they therefore repay their debts. Walter Wriston is also credited with a more perceptive remark, "The fact is that bankers are in the business of managing risk. Pure and simple, that is the business of banking."

11. Berger and Udell (2004) provide an interesting explanation for this by hypothesizing that institutional memory in banks erodes over time. Their "institutional memory hypothesis" asserts that loan officers gradually forget past lessons over the course of a lending cycle and hence ease credit standards when they should not. They also provide supporting empirical evidence.

Example of Successful Risk Management: U.S. Bancorp

This Minneapolis-based company, with $204 billion in assets at the end of 2005, had grown significantly through acquisitions until 2003. However, in 2003, USB's management and board made a commitment to the shareholders to return 80 percent of capital to investors. Since then, by focusing on its core markets, the bank has churned out a return on average equity of 22.61 percent, a return on average assets of 2.22 percent (for the four quarters spanning the last two quarters of 2004 and the first two quarters of 2005), and a Tier-1 capital ratio of 8.1 percent, despite intense competition that compressed its net-interest-margin to 3.99 percent.

How did U.S. Bancorp do it? Four factors were responsible. First, it significantly improved its credit-risk management, lowering its nonperforming loans to 0.46 percent of its loans, average for large banks. Second, the bank built a strong balance sheet and used it to compete aggressively for high-credit-quality corporate customers as well as obtain cash management, merchant processing, and other fee generating business with these customers. Third, the bank significantly improved its cost efficiency. And finally, U.S. Bancorp focused on developing a culture devoted to cost control and effective risk management.

Source: Engen, John, "Many Happy Returns," *Bank Director*, 4th quarter 2005, pp. 26–28.

Liquidity Risk: For some banks, Treasury bills, bankers' acceptances, Fed funds sold, demand balances, and short-term CDs of other banks are the primary source of liquidity. For many banks, however, liquidity is dependent on *liability management* (borrowing capability), which mostly involves borrowing in the interbank markets[12] and selling large negotiable CDs. Of course, the more acute the bank's need for such liquidity, the more costly/difficult it will be to acquire the desired liquidity. The 1984 failure of the Continental Illinois Bank and the 1998 failure of LTCM illustrate how liquidity can evaporate when most needed.

Liquidity management involves both routine (and therefore fairly predictable) cash needs and precipitous cash crunches. The former simply require putting a plan in place; they do not confront the bank with any major policy dilemmas.

Unpredictable cash needs are managed by keeping liquid reserves and by prearranging funding sources. Both measures are costly. The choice then is one of the right balance between the potential (but uncertain) benefit of reducing risk and the actual cost of doing so. The more astute bank treasures *diversify* funding sources for the very same reason that chief credit officers diversify earning assets. Such diversification can help to reduce the bank's dependence on one or a few credit sources that may dry up in a moment's notice.

12. Unsecured, short-term domestic interbank loans are called purchases of federal funds, referring to the fact that the money is borrowed by arranging for reserve balances at the Federal Reserve Banks to be transferred from the lending to the borrowing bank. Secured loans of fed funds are called repurchase agreement or "repos."

Banks use a variety of techniques to manage liquidity risk. Many regularly "test the waters" by raising more funds than they need at the time. Others use formal models to flag dangerously low or excessively high liquidity levels.

Of course, the Federal Reserve, with its discount window facility, is the ultimate provider of liquidity to a bank. The curtailed use of emergency credit under FDICIA means, however, that the discount window will be of less use in the future to banks truly in financial distress. FDICIA seeks to restrict discount-window use to solve liquidity problems rather than credit-risk problems.

Interest Rate Risk: In Chapter 4 we highlighted the attractiveness of interest-rate risk to banks when term premiums are high, and the losses that banks may be exposed to when they attempt to take advantage of these premiums. The potential losses due to interest-rate risk must be weighted against the possible gains from a duration-mismatched balance sheet.

Once the bank has settled on an appropriate procedure for measuring interest-rate risk, it has a variety of tools at its disposal to manage this risk. Chapter 8 discussed how the bank can use interest-rate swaps and futures to hedge its exposure. And Chapter 9 indicated the possibility of securitization as a way to cope with interest-rate risk.

It does not follow that banks should never accept and process interest-rate risk. The reward for brokering or even transforming interest-rate risk needs to be weighted against the exposure.[13] The discussion in the box below shows one bank's strategic approach to interest-rate risk management.[14]

Interest-Rate Risk Management by Citizens and Southern Corporation (C&S):[1]

Each year, C&S management used to set target financial ratios for earnings, credit risk, interest-rate risk, and other factors. In 1990, it decided to minimize interest-rate risk. By the end of February, the company claimed to have moved its gap to virtually zero.

C&S had a practice of financing itself primarily with core deposits, turning to purchased funds (jumbo CDs, fed funds, repos, and other short-term debt) only when core deposit growth proved insufficient. Thus, the duration of its liabilities was driven largely by the duration of its core deposits. Having computed this duration, C&S proceeded to shorten the duration of its assets. It finished 1989 with a 6-year average maturity in its asset portfolio, and in the first two months of 1990 it increased its holdings of CMOs with maturities ranging from 2 to 3.5 years. Like many banks, C&S valued the combination of high yields and high ratings offered by mortgage-backed securities, but it preferred CMOs to GNMA mortgage-backed securities because CMOs have shorter durations on average and hence lower interest-rate risk.

1. This bank was merged into Nations Bank, which later merged with Bank of America.

13. Eliminating interest-rate risk completely is like managing risk but not opportunity. Opportunity and risk are usually correlated. To manage risk and not opportunity is like driving a car with a brake, but no accelerator. To miss market opportunities is to invite competitors to fill the void. As Don Tomassino said in The Godfather III, "Your enemies always get strong on what you leave behind."

14. See also Holland (1990).

Off-Balance Sheet Risk: In assessing off-balance sheet exposure, the nominal value of the off-balance sheet items shown in the footnotes of balance sheets is not very informative. The *nature* of the risks must be carefully assessed. Extant computer models can estimate the values of contingent claims like loan commitments and standby L/Cs and provide some guidance in quantifying the bank's exposure. The bank also must carefully assess whether the exposure involves credit risk or interest-rate risk or both, and whether it jeopardizes its *future* liquidity.

Seven categories of risk have been defined for derivatives by U.S. bank regulators: counterparty credit risk, market risk, settlement risk, operating risk, liquidity risk, legal risk, and aggregation risk. Each of these is discussed in the box below.

Major Risks in Contingent Claims

- **Counterparty Credit Risk**: For contingent claims like interest-rate swaps involving a firm commitment, credit exposure is measured not by the notional amount of the contract but by the current cost of replacing it in the market This risk varies depending on the contingent claim. As mentioned in Chapter 8, this is really a form of interest-rate risk in the case of interest-rate swaps. The risk is somewhat different in the case of options. The buyer of the option faces the credit risk that the seller's financial condition may be such that he may not pay at the time of exercise. Working out the replacement cost exposure for most derivatives is a standard procedure.
- **Market Risk**: The market risk in contingent claims is similar to that of the underlying cash instrument. Contracts are used largely to hedge the price risks associated with the underlying cash instrument. There is, of course, a difference between contingent claims that are traded (such as options) and those that are not traded (such as loan commitments). Although there is no market price *per se* far a nontraded instrument, the price-related risk is higher for nontraded claims. Most banks divided market risk into components, such as interest-rate risk, exchange-rate risk, and commodity price risk.
- **Settlement Risk**: This is the risk that one of the parties to the contract may not honor it at the time of settlement. This is a form of credit risk.
- **Operating Risk**: Dealing in contingent claims requires knowledge of the mathematical tools of contingent-claims analysis and also high-tech information processing systems. Senior managers need to understand the activities of their "rocket scientists" and traders, and ensure that there is a shared sense of the bank's risk-management strategy. Operating risk can arise from inadequate internal controls and procedures in dealing with derivatives.
- **Liquidity Risk**: This is the risk that theoretically computed replacement values may be meaningless because the market has dried up just when a counterparty has defaulted. This is especially true of OTC markets and nontraded contingent claims.
- **Legal Risk**: The uncertainty associated with the legally binding obligations of various parties gives rise to this category. When is a contract legally binding? This is a particular difficult issue in international transactions, in part because of different national bankruptcy codes.

(*Continued*)

- **Aggregation Risk**: This risk, also known as interconnection or systemic risk, arises in contingent-claims transactions involving several markets and instruments. The risk is that the financial distress of one bank may jeopardize the entire transaction, possibly causing a major failure of settlement and payments systems.

Although banks have offered loan commitments and L/Cs for a long time, many have only recently begun to use derivatives. Thus, there may be a tendency to treat the management of off-balance sheet risk as distinct from its normal risk management. This should be avoided. What is important is the bank's *overall portfolio*, which includes on-balance sheet as well as off-balance sheet items. For example, when choosing its optimal degree of asset portfolio diversification on sport loans, the bank should take into account the possibility that some of its outstanding loan commitments will be exercised. How will each new loan created by a commitment takedown affect its loan portfolio variance? An analysis of possible future interest rates and takedown scenarios may prove helpful.

An integral part of managing off-balance sheet risks is measuring their magnitude. While there is no conceptual difficulty in doing this, not much data are available on the risk exposures of different banks. Surveys by the Bank for International Settlements show only the notional amounts outstanding. This is not an indicator of risk. Quite often banks estimate their off-balance sheet risk exposure no more than 3 to 4 percent of the notional value of the off-balance sheet activity.

U.S. regulators have concluded that the failure of a big institution as a result of contingent-claims activity is "unlikely in the current risk environment." Regulators also point out the danger of viewing contingent claims as inherently risky. It is possible, of course, that they actually reduce the net risk of the institutions that use them. In the final analysis, whether a contract is on- or off-balance sheet is purely an accounting issue and not an economic issue. Banks have often been misled by the accounting, partly because regulators have been misled too.

Management of Regulatory and Miscellaneous Risks

Regulatory Risk: Regulatory restrictions on banking activities are aimed partly at preventing banks from taking excessive risks. But these restrictions and proscriptions may lead to an increase in risk because they create perverse incentives. For example, precluding the holding of equity in all firms may lead banks to hold very risky debt in high-risk firms. This is similar to the effect of cash-asset reserve requirements (recall Chapter 11).

Another reason why regulatory restrictions may inadvertently increase risk is that they inhibit banks from innovating and adapting to new technology and evolving customer needs. They do this in two ways. First, the restrictions limit the range of options the banks can exploit in a constantly changing environment, and thus lead to higher risk. Second, as noted earlier, these restrictions tend to experience unpredictable changes through time, and since this unpredictability creates a nondiversifiable risk for the bank, its cost of capital goes up. A consequence is that banks may be put at a competitive disadvantage relative to their less-regulated nonbank competitors.

In a survey of selected banks, change was cited as the biggest component of the regulatory burden.[15] Apart from the risk associated with unpredictable regulatory change, there are significant out-of-pocket costs for banks when a regulation is changed. These include the costs of retraining employees, tearing up old forms, computer reprogramming, and consulting with legal and compliance experts on how best to adapt to the change.[16]

To understand the reasons for regulatory unpredictability, it is useful to recall the distinction between *nondiscretionary* and *discretionary* regulation (Chapter 11). Nondiscretionary regulation refers to regulatory *rules*, such as cash-asset reserves and capital requirements. Discretionary regulation involves regulatory *judgment*, as in the case of bank closures.

Whereas nondiscretionary regulation may change because of unpredictably evolving environmental conditions, randomness in discretionary regulation may represent an instrument of regulatory policy. Many discretionary regulations are *purposely* kept *ambiguous*, and this ambiguity creates another source of risk for the regulated firms. For example, the *Community Reinvestment Act* (CRA) "encourages" banks to invest in low- and moderate-income neighborhoods in their lending areas, but it does not stipulate what percentage of lending must be done in those neighborhoods nor does it precisely identify which borrowers qualify as low-income borrowers in the bank's lending area. Yet, when a bank is found in violation of the CRA guidelines, the penalties can be severe, ranging from a directive to improve its CRA performance to the denial of an application to open a new branch or merge with another bank. The ambiguity in the CRA is somewhat paradoxically justified as a desire to not dictate credit allocation to banks. Similarly, the government's "Too Big To Fail" (TBTF) policy (see Chapter 10) is ambiguous because it does not tell the banks which of them will be covered by its policy and which will be allowed to fail.[17] This policy is often justified on the grounds that such ambiguity can be effective in deterring moral hazards.[18]

Whatever the rationale for the ambiguity, it is pervasive in government regulation. Such ambiguity also increases the bank's risk and its cost of capital, and undermines competitiveness. For example, suppose a bank plans to open three new branches and makes asset portfolio adjustments in view of the planned branch openings. If it has a good CRA rating, the bank may believe that regulatory approval of the new branches will be routine. When knowledge of its application becomes public, however, local community groups protest on the grounds that the bank has not met the community's credit needs adequately. Consequently, regulatory approval may be delayed or denied, imposing heavy costs on the bank. The key is that regulators often fail to internalize the cost of their unpredictability. This cost is borne by the bank's shareholders and possibly by its customers.

Besides these somewhat unpredictable components of regulatory risk, there are also components that are more clearly defined and easier to predict. Banking

15. See Carroll, Danforth, Golembe, and Laub (1989).

16. A study by Barefoot, Thakor, and Beltz (1993) estimated that the average cost for banks of complying with just consumer protection regulation alone (for example, CRA, HMDA, BSA, and RESPA) was about 18.3 percent for banks of (after-tax) net income, and that this percentage was greater for smaller banks.

17. Even though the OCC officially announced a list of banks that were "too big to fail," the government subsequently applied the TBTF policy to banks that were not on the comptroller's list.

18. See, for example, Boot and Thakor (1992).

agencies regularly dispatch examiners who scrutinize banks and submit detailed reports. While the recent spate of bank and S&L failures highlighted some of the weaknesses in the examination process, bank examiners are "troubleshooters" who spot problems for which they instruct bank management to take corrective action. In assessing a bank's soundness, examiners consider the bank's exposure to credit, liquidity, and interest-rate risks, the bank management's ability to control these risks, and the bank's financial resources (capital and liquidity) to cope with these risks. The bank is then given an overall rating, called the CAMELS (Capital, Assets, Management, Earnings, Liquidity, and Sensitivity to Market Risk) rating (recall Chapter 11).

Examiners also look at financial ratios and use a comprehensive ratio analysis system, "The Uniform Bank Performance Report," to compare the bank's financial ratios with those of its peer group. In addition, each regulatory agency has its own "early warning" system, designed to detect emerging problems. During their on-site visits, examiners are expected to investigate problems identified by these computer monitoring systems.

If the bank's management does a poor job of anticipating problems that the examiners might point out, on-site visits by examiners can cause serious disruptions for the bank, forcing it to make numerous unplanned changes at short notice. The key to minimizing such disruptions is to develop effective in-house monitoring systems to anticipate the concerns of examiners and then to act on them in a planned and orderly manner prior to examinations. The resources the bank should optimally devote to regulatory compliance will depend, to a large extent, on the bank's size, and the nature of its operations.[19]

It is virtually impossible for a bank to totally eliminate regulatory risk. But effective intrafirm communication systems can be designed so that all parties concerned, including the board of directors, are kept apprised of the bank's policies and the steps it is taking to satisfy regulatory requirements. Of course, it may not be optimal for the bank to aim for complete regulatory satisfaction. The interesting question is: How do you manage when complete regulatory compliance is not even possible?

(i) Electronics Funds Transfer (*EFT*) *Risk*: Billions of dollars are transferred electronically each day among banks in the United States and in other countries, and the volume of interbank transfers has been growing rapidly, putting enormous strains on the system. Some observers believe that payments system risk is the greatest of the risks facing banks.[20] To understand the risks in funds transfers, first we must understand the process by which funds are transferred. In the box below, we describe this process and the major clearing mechanisms, after which we discuss EFT risks. You can skip the box if you are familiar with these details.[21]

19. The current wisdom is that once a bank reaches $100 million in assets, it should hire a full-time compliance officer.

20. *The Economist* (1993) quotes *Sir* Dennis Weatherstone, chairman of J.P Morgan, as stating that settlement risks and daylight exposure worry him more than off-balance sheet risks. See "International Banking: A Comedy of Errors," *The Economist* (1993).

21. The discussion below is based on Humphrey (1989); Mengle (1992), and Mengle, Humphrey, and Summers (1987).

Clearing and Settlement Systems for Funds

Clearing refers to processing a trade involving a financial security and establishing the obligations of the parties involved. *Settlement* refers to the transfer of value between the parties so that the trade is consummated. There are two basic steps in the clearing and settlement process. The first is to convey the details of the trade from traders to the back office of the bank. The second is to check with the buyer and the seller to ensure that both agree on what is to be traded and on what terms. This minimizes delivery problems.

The Role of Banks: Banks and the interbank payment system are at the heart of the clearing and settlement mechanism for the money market. Banks connect the participants in the money market by acting in three capacities: (i) as agents for issuers of money market instruments, they issue and redeem instruments in the market and maintain registration records; (ii) as custodians of instruments, they provide a safekeeping service to investors; and (iii) as clearing agents, they transfer securities and payments for these securities across transacting parties.

Transfers between banks take place at the hub of the money market, *the interbank payment system*. Even when instruments are traded outside the banking system, payment takes place through banks. The payment system links banks to each other, and almost all interbank payments now occur electronically over *wholesale wire transfer network*.[1] *Fedwire*: *The main wholesale wire transfer network in the United States Fedwire*, which operates through bank reserve accounts at the 12 Federal Reserve Banks. During 1991, about 260,000 Fedwire transfers totaling about $766 billion occurred on an average day.

Bank of Tinseltown

Assets	Liabilities
Reserves, −$50 million	

Federal Reserve Bank

Assets	Liabilities
	Reserves, Tinseltown −$50 million Reserves, United −$50 million

United Bank

Assets	Liabilities
Reserves +$50 million	

FIGURE 13.4 A Fedwire Transaction

Figure 13.4 shows how Fedwire is used to complete a fed funds transaction. Suppose the Bank of Tinseltown has $50 million in excess reserves, whereas the United Bank is $50 million short of required reserves. A broker arranges a fed funds transaction between these two banks. Settlement occurs through the Federal Reserve Banks in the districts to which these banks belong. Tinseltown's reserve account at its Federal Reserve Bank is drawn down by $50 million, and immediately after that United's account with its Federal Reserve Bank is increased by $50 million. Fedwire payments are final and irrevocable. As far as the Fed is concerned, this transaction simply moves reserves from the account of one bank to the other. The next day, United uses Fedwire to repay Tinseltown, thereby reversing the earlier transaction.

(*Continued*)

A key feature of Fedwire transfers is that the settlement is on a bilateral, trade-for-trade basis, also known as *gross settlement*. An alternative settlement procedure would be to consolidate transfers into net positions either between banks or between the network and banks. Such a procedure reduces the number of interbank transfers that occur, and the system would be called a netting system.

Netting can be either *bilateral* or *multilateral*. Bilateral netting combines gross obligations between bank into net obligations so that each pair of banks in the system exchanges only one settlement payment. Multilateral netting combines each bank's bilateral net positions into "net net" obligations between the bank and other banks in the system (that is, the network). The network adds up the amounts owed to and owed by each bank to compute the "net net" for each bank; this "net net" is essentially the bank's net position versus the network. Each bank is either a net creditor or a net debtor. Moving to bilateral netting and then to multilateral netting can mean substantial reductions in the number of actual exchanges between the participants. For example, suppose there are m banks in the system and there are two transactions between *each* of these banks, that is, bank 1 has two transactions with each of the other m – 1 banks, say one in which it is a debtor and the other in which it is a creditor. Then, the table below gives the total number of transactions with each settlement procedure.

TABLE 13.1 The Relationship Between Settlement Procedure and Transaction Volume

Type of Settlement	Total Number of Transactions
Gross Settlement	$m \times (m - 1)$
Bilateral Netting	$\frac{m(m-1)}{2}$
Multilateral Netting	m

To illustrate, suppose there are six banks, each involved in two transactions with the other banks. Then, the total number of transactions is 30 with gross settlement, 15 with bilateral netting, and six with multilateral netting.

CHIPS: The Clearing House Interbank Payments System (CHIPS) is a multilateral netting system. It is owned and operated by a private organization called the New York Clearing House. CHIPS transfers funds but not securities. It is used largely in connection with international transactions such as Eurodollars and foreign exchange. It handles 95 percent of all U.S. dollar cross-border payments. In addition, payments include financial settlements (e.g. loan and interest payments), commercial payments, and off-shore investments.

In 2006, CHIP processed on average 285,000 payments each day with a gross value of \$1.4 billion. The organization of CHIPS is hierarchial. A subset of the participating banks (20 out of 126) settle directly with CHIPS while the others are required to settle on the books of one of the settling banks. *Figure 13.5* shows how settlement takes place through CHIPS. The special CHIPS net settlement account starts out with a zero balance prior to each settlement, and ends with a zero balance when the settlement is complete. It is used for nothing else.

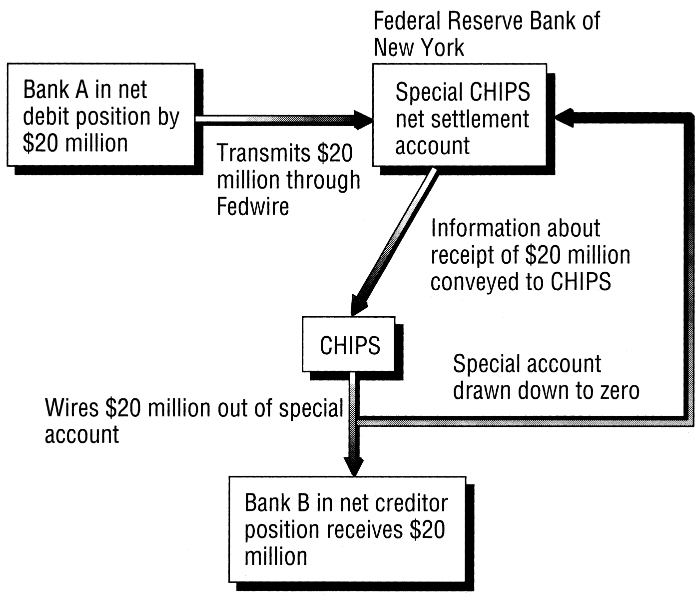


FIGURE 13.5 Settlement Through CHIPS

SWIFT: The Society for Worldwide Interbank Financial Telecommunication (SWIFT) is a nonprofit cooperative chartered under Belgian law and owned by 1,885 participating institutions in 73 countries, including the United States. Unlike Fedwire or CHIPS, SWIFT is not a funds transfer system. Instead, SWIFT payment messages instruct banks to transfer funds by means of accounts at correspondent banks. For example, suppose the Bank of Tinseltown is a correspondent bank for both the United Bank and the Bank of Amsterdam, and also suppose that the Bank of Amsterdam wishes to make a payment to the United Bank. This payment could be made by Amsterdam sending a SWIFT message instructing Tinseltown to reduce Amsterdam's correspondent account and to increase United's by the amount of the payment. If United and Amsterdam do not use a common correspondent bank, then two correspondent banks would be involved and Amsterdam's SWIFT message would instruct Tinseltown (its own correspondent bank) to transfer funds to United's correspondent bank using an interbank network like Fedwire or CHIPS.

1. Wholesale wire transfer networks link banks with each other. In contrast, retail wire transfer systems, such as on-line automated teller machines (ATMs), link banks with customers.

There are various types of EFT risks. Policy discussions distinguish four major risk categories,[22] which are discussed below.

Credit risk arises in EFT because a bank may send funds electronically for a customer before that customer has sufficient funds to cover the transaction. The bank has effectively made a loan to the customer, and as in any other credit

22. See Parkinson et al. (1992), for example.

transaction, there is a possibility that the bank may not be repaid. The risk is normally low because the loans have extremely short maturities, but it is present nonetheless. To limit this risk, banks often place limits (caps) on daylight overdraft credit extended to particular customers, with on-line approval required for large transactions.

Credit risk may or may not arise for the bank that receives funds. If the funds are sent over Fedwire, there is no risk to the receiving bank because the funds are credited to the receiving bank's reserve account without recourse, and the Fed guarantees the transaction. That is, in order for the Fedwire transaction to take place, it is not necessary for the sending bank to have sufficient funds in its reserve account to cover the transaction. For example, in *Figure 13.4*, if the Bank of Tinseltown has only $40 million on deposit as reserves, Tinseltown incurs a "daylight overdraft" (it has taken a loan with a maturity of up to a day) of $10 million. The reserve account is thus allowed to go negative during the day as long as the deficit is made up before the close of business. Since the receiving bank received final payment at the time of the transfer regardless of whether the overdraft is ultimately covered, the Federal Reserve Bank bears the credit risk arising from the possibility that Tinseltown will default. However this "socialization" of credit risk by the Fed means that systemic risk has been eliminated with Fedwire. On the other hand, if the funds are sent over a privately owned system, there is a risk for the receiving bank. Under CHIPS, all transactions are considered conditional until the end of the day. To see how this creates credit risk, suppose that the United Bank has received a credit for $10 million in the name of its customer, Motown Car Company. The Bank of Amsterdam has remitted this amount through CHIPS on behalf of one of its customers. Before the end of the day, Motown withdraws the entire amount. However, if Amsterdam fails to meet its net settlement obligation at the end of the day, United will not actually receive the $10 million owed to it. Although it has recourse to Motown's account, there is nothing left in the account, so it has effectively extended Motown a $10 million loan.

(ii) Liquidity Risk: This risk arises from the possibility that settlement could be delayed because of temporary unavailability of funds.

(iii) Systemic Risk: The liquidity and credit risks associated with CHIPS could end up being *systemic* due to possible contagion effects. In the context of the earlier example of Amsterdam and United, the failure of Amsterdam to settle a particularly large net debit position could cause a chain reaction of settlement failures among other participants, some of which might depend on the receipt of payments from Amsterdam in order to fund their own obligations. These ripple effects could lead to a systemwide failure.

(iv) Operation Risk: This risk arises from the possible breakdown of information-processing technology or other operational elements in the clearing and settling mechanism. An example was the computer problem at the Bank of New York in 1985 that allowed the bank to accept securities but not to process them for delivery to dealers, brokers, and other market participants. Consequently, the bank's reserve account was debited for the amount of the securities, but the bank was unable to resend them and collect payment. This led to a growing daylight overdraft in the Bank of New York's reserve account. Since the problem could not be fixed by the end of the business day, the bank was compelled to borrow from the discount

window; the problem was fixed overnight and the discount-window loan was repaid the next day.

We turn next to how these risks can be controlled. There are seven *risk-control* measures, which are discussed below.

(a) Membership Standards: These help to exclude participants who lack the financial strength and operational expertise to ensure that settlement obligations can be met.

(b) Quantitative Limits on Risk Exposure: These limits include net debit caps and bilateral net credit limits. Net debit caps are limits on the size of a bank's combined daylight overdraft on Fedwire and net debit position on CHIPS. That is, these are limits on how much a bank can owe over the wire transfer networks. Bilateral net credit limits specify the maximum net transfer a bank on CHIPS is willing to receive from a particular sending bank. These limits permit a bank to control its own exposure to other banks. Both these quantitative risk-exposure limits were part of the Federal Reserve risk-control policy adopted in 1986.

(c) Collateral: Participants can be asked to put up collateral to cover their obligations to the system.

(d) Explicit Pricing of Daylight Overdrafts: As explained earlier, a daylight overdraft is an intraday negative reserve account balance on Fedwire and an uncovered net debit position on CHIPS. These overdrafts can last anywhere from a few minutes to most of the day. Until recently, banks were not charged interest for intraday credit. There were neither any controls nor any costs associated with the use of daylight overdrafts. Banks naturally came to rely on the free credit provided by daylight overdrafts rather than developing alternative payment arrangements. To control the Fed's credit risk from daylight overdrafts with Fedwire, the Federal Reserve adopted an explicit pricing policy for intraday credit. In 1992 the Federal Reserve approved a charge on daylight overdrafts that exceed 10 percent of an institution's risk-based capital. The idea is that pricing should impose an explicit cost on the use of intraday credit and thereby provide incentives for a more efficient allocation of that credit.

(e) Loss-Sharing Arrangements: Under a loss-sharing agreement, member banks in the system share the losses caused by the failure of another member to settle. For such an agreement to work, two features are needed. The first is *settlement finality*, which means an assurance that settlement entries will not be reversed in the event of one bank's failure to settle. The other is *collateral*, that is, for the credibility of the loss-sharing agreement, member banks should contribute collateral to a clearing fund; this can be drawn down in the event of a settlement failure. This collateral also provides the necessary incentives for banks to monitor each other (recall Chapter 5). CHIPS adopted settlement finality and a loss-sharing agreement in 1990.

(f) Obligation Netting: This means combining offsetting gross payment of securities obligations into net obligations. As discussed previously, netting can be bilateral or multilateral and can reduce operational risks by reducing the volume of transactions that pass through the clearing and settlement system.[23]

23. The *Government Securities Clearing Corporation* (GSCC) was established in 1986 to provide netting of government securities trades for banks and other securities brokers and dealers. The multilateral netting is similar to that under CHIPS, except that the numbers would refer to sales or purchases of a specific issue of government securities instead of CHIPS funds transfers.

(g) Book Entry and Delivery-Versus-Payment: Book entry eliminates the physical delivery of instruments, whereas delivery-versus-payment means that the exchange of funds and securities is almost simultaneous. Book entry reduces operational risks and transactions costs, whereas delivery-versus-payment reduces credit risk.

Risk of Fraud: Fraud risk is second only to credit risk as a cause of bank failure. It was estimated that in the 1950s and the 1960s, two-thirds to three-fourths of bank failures involved fraud, but this fraction fell as failures increased. The overall implications of fraud extend well beyond the direct losses to the bank. When management engages in fraudulent behavior, it also signals a lack of commitment to the business and carelessness in dealing with the risks faced by the bank. Employee fraud and mismanagement often go hand in hand.

Fraud need not only be internal. It can also be perpetrated by outsiders, or by a combination of insiders and outsiders. External fraud can take the form of check kiting, counterfeiting, falsified loan requests, and credit card abuse.

Internal fraud is often hard to detect because the schemes are elaborate and perpetrated by people the bank trusts. Internal fraud encompasses everything from extravagant consumption of perks to illegal loans to management. Inappropriate loans to insiders and to their families and friends have led to numerous conspicuous bank failures, for example, the U.S. National Bank of San Diego, Franklin National Bank, the United American Bank of Knoxville, and the Penn Square Bank. Indeed, when a bank is in financial distress, incentives for fraud become stronger, just as the incentives for excessive risk-taking increase when capital is low.

One difficulty in dealing with fraud is that whenever it is encountered, internal audit is almost always neutralized. This suggests that there is a degree of *optimal redundancy* in auditing. The bank should search for the best combination of internal auditing, external auditing, examinations, and legal division checks.

Miscellaneous Risks: Most of the risks discussed thus far are integral to the business of banking. But banks also face a variety of risks that have little to do with banking *per se*. These include the risk of losses due to fire, earthquakes, floods, computer breakdowns, and so on. The bank can be presumed to have no special expertise in processing such risks. Thus, the most efficient way to manage these miscellaneous risks is to seek *external insurance*.

Another miscellaneous risk is that key employees may leave the bank. Although the bank can purchase "key-employee" insurance against this risk, large firms usually self-insure.

Management of Routine Opportunities

In managing risks and coping with change, a bank may pay inadequate attention to the management of its existing customers and markets. Since these currently sustain the bank, management should be aware of possible threats to the bank's position in these markets. Existing customers are assets of the bank that often increase in value with relationship longevity. Consequently, these relationships need to be managed like the other assets of the bank with monitoring and marketing efforts. Quite often, the bank's customers will receive tempting offers from other banks that wish to entice them away. If the bank does not respond appropriately, the customers may be lost.

The management of routine opportunities requires an integrated approach that assigns a group to deal with all aspects of the business of each particular class of customers. Organizationally, the banks that are best suited to implement this approach practice *relationship banking*. In this case, the customer deals with the same group within the bank for everything from business solicitation to credit analysis to the determination of credit terms. This reduces the likelihood that existing customers and markets will receive inadequate attention.

An appropriate organization structure should be used in concert with an effective internal monitoring scheme to ensure that current markets are being adequately served. This means adjusting prices as supply and demand conditions change, and in some instances, lowering prices *before* competitors enter the fray. The key for the bank is to understand the dynamics of each market and the bank's changing role.

Crisis Management and Enterprise Risk Management

The six most frequently encountered crises in banks are: (1) bad loans (see Chapters 5 and 6), (2) fraud, embezzlement, and white collar crime (see Chapter 10), (3) hostile takeovers, (4) unexpected executive succession, (5) problems with regulators, and (6) image problems, rumors, bad press, and lack of public confidence.[24] In practice, the level at which the crisis is handled depends on the nature of the crisis. Crises arising from physical causes (for example, computer failures and fires) tend to be handled at lower levels of management than those arising from nonphysical sources (for example, hostile takeovers and image problems, and so on). One reason for this pattern of delegation is the magnitude of physical versus nonphysical crises. Physical crises tend to be localized, and are best handled by the employees closest to the event. By contrast, an intangible crisis such as a hostile takeover affects the *entire* bank. Only top management has the perspective to respond effectively to such organization-wide threats.

In *Figure 13.6* we have sketched the process by which crises arise and are managed. The first stage after the occurrence of a crisis is *problem sensing*. At this stage, management interprets the crisis situation and its dimensions.

The subsequent crisis management stage is shown as circular and interconnected to signify that crisis management is a nonsequential, interactive process.[25] In addition to problem sensing, crisis management involves the four activities shown in *Figure 13.6*: (1) *decision response*, which refers to the act of choosing a course of action; (2) *resource mobilization and implementation*, which refers to the act of mobilizing resources to implement the decision response; (3) *internal information flow*, which refers to the flow of information among management and employees within the firm; and (4) *external information flow*, which refers to the flow of information to the shareholders, depositors, borrowers, regulators, and the public.

Quite often, crisis resolution involves the displacement of bank employees, particularly if the crisis can be traced to the misbehavior of these employees. And the more serious the crisis, the higher is the likely level in the organization at which employees are affected. A major crisis that can threaten the very survival of an institution is from an event that damages its reputation. Banker's Trust realized

24. See Reilly (1988) for further details.
25. See Mitroff, Shrivastava, and Udwadia (1987).

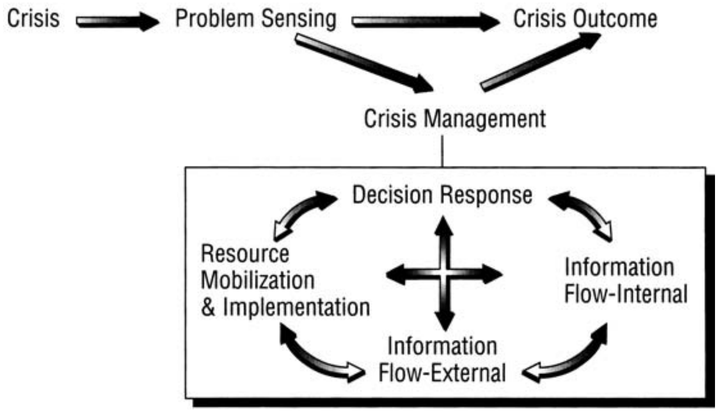


FIGURE 13.6 The Process of Crisis
Source: Mitroff, Shrivastava, and Udwadia (1987).

this from its experience with swap transactions with P&G and Gibson Greetings, as a result of which it was embroiled in litigation. Arthur Andersen imploded over the accounting scandal that brought down Enron in 2000.[26] The lesson from these scandals is that intermediaries whose *raison dé tre* is their reputation and their credibility in providing certification services—like auditing firms and many types of financial intermediaries—cannot afford reputational setbacks. The risk of such setbacks is a major risk that must be recognized and managed.

What kinds of banks are best prepared to manage crises? It has been found that the crisis readiness in a bank is positively related to its *long-term strategic orientation* (that is, its focus on long-range planning) and negatively related to *organizational complexity* (the number of activities performed by the organization, the number of branches, and so on). This seems to suggest that as banks engage in a greater variety of activities, they will have to make a special effort to improve their crisis readiness.

Recently, the idea that risks should be comprehensively managed has been put forth under the label of Enterprise Risk Management (ERM). Basically, ERM is a structured and disciplined approach to managing risk by aligning the organization's strategies, processes, technology and knowledge on an enterprise-wide basis, cutting across functional and departmental boundaries. Good information and effective processes are the key to ERM, and good ERM also equips the organization more effectively to sense a crisis before it occurs.

Strategic Planning

As information processing technology becomes more sophisticated and markets become more integrated, the operating environment for banks is increasingly complex and competitive. Managing in the context of a long-term strategy therefore becomes more important. Regulatory subsidies and exclusive franchises led to

26. An interesting study by Chancy and Philipich (2002) found that the Enron scandal affected most of Anderson's other auditing clients adversely in terms of negative stock returns on the announcement of the scandal, with the Houston office clients being affected more adversely.

complacency and resistance to change on the part of bankers.[27] In the contemporary banking environment, strategic options for banks are likely to be quite limited within traditional banking segments, so that seeking new products and markets, within the constraints imposed by regulation, will be very important. Successful banks will do well in the following areas: innovation, cost accounting and pricing, management of risks, management of human resources, marketing, management of acquisitions, planning, and organizational design. We discuss each briefly.

- **Innovation:** How nimble a bank is in adopting new products and services will be an important determinant of a bank's ability to continue to be profitable. In a competitive environment, profit opportunities are ephemeral. As the investment banking industry illustrates, when a new product is introduced, the firm that introduces a new product is a monopolist and can earn monopoly rents until its competitors, encouraged by its profits, decide to enter the market. Once a sufficient number of entrants are in the market, the monopoly rents disappear. Thus, the incentive to innovate depends on the costs of innovation and how long it takes before competitors copy the innovation and compete away monopoly rents. Financial services are easy to copy, so rents are quickly dissipated. To consistently earn rents, firms must be able to "roll out" new services quickly with minimal dislocations. Various forms of structured finance, securitization, and the proliferation of off-balance sheet contingent claims are all examples of successful financial innovations that have generated rents for banks. In the box below we discuss the economics of innovation.

The Economics of Innovation

What forces stimulate innovation, and what forms does innovation take? To answer these questions, we define innovation first.[1] Financial innovation is a process by which the financial system adapts to environmental shocks such as changes in technology, demand, and public regulation. This adaptation is manifested in the mix of claims produced, the delivery systems through which the claims are marketed, and the organizational forms and industrial structure of the financial services industry. When external shocks are moderate relative to the adaptive capability of the financial system, there is a smooth passage from one state to the next. But when shocks are large or when the financial system's adaptive capability is impaired, disturbances result in financial crises, recessions, and related symptoms.

This adaptive process is illustrated by the American experience in the 1970s during which massive real resources were diverted to financial applications such as cash management. This was because inflationary expectations ratcheted upward. Credit contracts were reconfigured to shorten durations because depository institutions lost core deposits. These changes were facilitated by technological advances in data processing and telecommunications.

To understand the forms that innovation can take, we can view the financial system as being composed of claims, delivery systems, and organizations. A claim is simply a financial contract that embodies a mix of attributes. As this mix is altered the

(*Continued*)

27. The discussion in this section is based in part on Parker (1981).

claim is redefined. A *delivery system* is a financial asset used to market claims. Delivery systems tend to be market-specific rather than product- or service-specific. Retail commercial banking has traditionally employed a brick and mortar delivery system that is becoming obsolete. Life insurance companies and commercial lending have relied mostly on labor-intensive personalized selling. Mutual funds use phone, mail, and wire distribution systems. Credit and debit cards represent yet another distribution technology. *Organizational form*, the third facet of the financial system, is simply the structure of the organization that delivers claims. It can be viewed as the outcome of the tension between the technology of production and public regulation. That is, regulation may preclude adopting the most technologically efficient organizational form.

The interaction of regulation, information technology, and changes in the operating environment of banks has affected all three facets of the financial system. Financial claims have proliferated in recent times. Many view this proliferation as a move toward making the market more complete (recall our discussion of market completeness in Chapter 1). While this market completeness has been facilitated by advances in information technology, it is also useful to note that regulation plays an important role because regulatory restrictions increase market incompleteness by precluding the offering of particular claims or the availability of particular distribution systems. In other instances, regulation (for example, reserve requirements) adds to the cost of producing specific claims. Whatever the effect incentives are generated to produce alternative financial claims.

Regulatory restrictions on organizational forms, particularly spatial and functional limitations, have constrained the types of delivery systems. Without these restrictions, the technology of brick and mortar would probably have led to extensive branching systems such as those in the United Kingdom and Canada. Similarly, real estate securities distribution, accounting and legal services, and insurance might have been distributed through the same brick and mortar delivery system were it not for legal restrictions. Relaxation of public regulation of organizational forms seems to be a major contemporary theme. Consequently, we have observed significant innovations in organizational forms and delivery systems, leading to franchising, networking, and shared ATM systems. Other examples are the interstate expansion of thrifts and bank loan production offices (LPOs), mutual fund sales by banks, and the entry of both thrifts and banks into discount brokerage.

The continuation of this trend toward increased scale and scope of delivery systems probably will lead to a financial services industry organized in tiers of firms. There will be a small number of worldwide systems at the top followed by nationwide delivery systems that will serve mass markets without brick and mortar or labor-intensive technologies. These nationwide systems will be complemented by a larger number of specialized distribution systems for narrow segments of the market. Routine financial services will be delivered electronically, but commercial lending, underwriting, and many forms of insurance will continue to require more labor-intensive distribution methods. Powerful forces will continue to push toward a coalescence of commercial banks, insurance companies and pension fund managers. The amount of consolidation will depend in part on the scale and scope economies that are available.

1. This discussion is based in part on Greenbaum and Higgins (1983), Boot and Thakor (1997).

Of course, innovation is inherently risky. Occasional failures should be accepted as part of the routine risks of innovating. Regulatory uncertainty, and the resulting risk for banks, can increase the bank's cost of capital and retard innovation. In general, anything that increases the bank's cost of capital will reduce innovation.

- **Cost Accounting and Pricing:** Banks have not been very astute with respect to their accounting practices and in their pricing. Some observers have commented sardonically that if borrowers used the same systems that banks have used, banks would deny them credit.

  One way in which banks have been inefficient is in the allocation of costs to various units. Banks have traditionally "bundled" their services for pricing purposes, resulting in cross-product subsidization. With increasing competition, it becomes necessary to unbundle services so that each becomes "a tub on its own bottom." The scope for cross-subsidization is therefore eliminated and each service must be self-sustaining.
- **Management of Risks in Strategic Options:** Since a strategic plan involves planning for long-horizon possibilities, the risk involved in choosing strategic options is likely to be greater than in routine operations. Procedures for estimating these risks and assessing the values of different strategic options must be developed separately from those used to measure the values of routine options. This is partly because the *time horizon* over which the value of a strategic option is evaluated is important. In many instances, early losses may have to be incurred in order to gain the desired market penetration for subsequent profits. Unless the time horizon is defined to be sufficiently long, the strategic option may not look profitable.

  A separate assessment of strategic options does not mean that the *interaction* between the risks of the strategic options and those due to routine operations should be ignored. When a long-term plan is adopted, the bank also must put in place mechanisms to hedge the interest-rate risk and other risks associated with the plan. The compatibility of these mechanisms with existing measures to cope with routine risks must be considered (recall our discussions in Chapters 4–9).
- **Human Resources Management:** Heightened competition and diminished governmental subsidies have elevated the importance of managerial talent in banking. Human resources management (HRM) has three key aspects. First, it involves hiring the best managerial talent, given the organization structure and resources. Second, designing efficient communication and employee incentive schemes can help to maximize productivity.[28] Given the pervasiveness of principal-agent problems, aligning the incentives of employees with those of shareholders is critical. This alignment is best achieved in a cooperative and collegial environment. Communication training and adopting the proper leadership values are important ingredients in producing such an environment. Third, HRM calls for the optimal mix of human and technological resources, so as to eliminate excess capacity and waste. Banking is, after all, an information and transactions processing business; its key elements are human resources, financial capital, and information processing technology.
- **Marketing:** When banks owned their markets, they did little advertising and acted mostly like novices in marketing. Market segmentation and product differentiation

28. The importance of incentive schemes in the financial services industry was highlighted by the experience of Continental where incentive contracts for loan officers rewarded asset growth rather than net profits, and later by the Salomon Brothers Treasury bills scandal in 1991 when the incentive contracts for managing directors were blamed for the company's illegal bidding behavior.

were rare. Increased pressure on profit margins has made these marketing issues much more important now.[29]

Banks' traditional sources of profits are drying up. Lending to Fortune-1000 corporations, once the big banks' lucrative *raison d'etre*, has dwindled. Credit card markets are under siege from powerful new competitors. These changes have compelled many banks to specialize. Now mega (money-center) banks focus on retail lending and trading. Super-regionals focus on retail and middle-market lending, and are heavy users of technology. Community banks are involved in regional and small-business lending and are more labor intensive. And then there are high-net-worth banks that specialize in "upscale" retail lending and financial services. The basic idea is that age, geography, and consumer wealth create numerous submarkets, so that banks can specialize and capture more of the "consumers' surplus" by fine-tuning the product to best meet a particular group's preferences.

Advances in information technology have facilitated market segmentation. Computers enable more banks to analyze their customers by age, income, and geography. They are also better able to reach those segments in the wake of deregulation, because they can offer products and interest rates that they could not offer in the past. Since this trend is likely to continue, strategic management should assign a significant role to marketing.

The importance of marketing was driven home for banks when they experienced a backlash during the 1990s after they automated services by leveraging the Internet and ATMs. Many banks closed branches and imposed teller fees, anticipating that electronic banking would sweep through the industry. Many discovered, much to their chagrin, that many of their customers preferred person-to-person contact.

Eager to win over customers and acquire their accounts and financial service business, banks are now increasingly turning to the latest marketing techniques and trying to learn from successful marketing practices *outside* the banking industry. There is good reason for this. A study by CIBC World Markets showed that at least 700 people walk into Starbucks daily, whereas the average bank customer visits a branch just 2.9 times per month, according to Synergistics Research Corporation.[30] Banks are therefore now copying the marketing approaches of successful retailers like the Ritz-Carlton.

Commerce Bancorp is one of the first to adopt the retail mindset and model its branches after successful retailers like Home Depot, Victoria's Secret, and Target. The New Jersey-based bank offers free coin-counting machines in all its branches, 7-days-a-week banking, free personal checking accounts, and personal greeters. The bank also uses mystery shoppers to gauge how well it is performing.

Another bank that provides an example of good marketing is Umpqua Bank, a subsidiary of Umpqua Holdings Corporation. Good customer service is part of its corporate culture. Every staff member is trained in all areas of the banks, so that a customer asking a bank employee for help is never confronted with a response, "it's not my job." Each branch is designed to create a neighborhood gathering feel. The bank serves its own brand of coffee, offers newspapers and periodicals and allows

29. Sease and Guenther (1990) quoted Thomas Hanley, a bank analyst then at Salomon Brothers (and now at First Boston), Inc., as saying, "... where does the profitability of the (banking) industry come from? It's not a pretty picture."

30. See Bergman (2005).

customers to log in at the Internet Café to check e-mails, surf the net or download songs on to a customized Umpqua CD. It even hands out chocolates with every receipt. By 2005, more than 20 financial institutions from around the world had visited the bank to study its culture and customer service policies.

- **Management or Acquisitions:** Heightened competition and more relaxed regulation have resulted in consolidation in the banking industry (see Chapter 14). This consolidation has been achieved mainly through friendly mergers and acquisitions rather than through hostile takeovers, although there have recently been some hostile takeover attempts. The principal motivating factors for these mergers are economies of scale to be exploited by merging, and the fact that many banks are in the process of redefining their markets. As mentioned earlier, large money center and super-regional banks have been losing the "Fortune 1000" firms to the capital market; commercial paper issues by U.S. corporations grew sixfold during 1975–95. So these banks, particularly the super-regionals, are increasingly turning to middle-market borrowers who have traditionally been served by regional banks. Rather than compete directly with regional banks for these borrowers, the super-regional banks have decided to gain access to these customer bases by acquiring regional banks; most of the consolidation has been among regionals and super-regionals, although there have been some huge within-market consolidations (for example, Bank of America and Security Pacific).

  For a bank that is involved in an acquisitions program, key management issues include the following. First, *organizational compatibility* must be considered. How compatible are the corporate cultures of the acquiring and target banks? How different are the credit policies and the decision-making systems, and how costly/disruptive is it likely to be to harmoniously blend these different cultures? Second, issues of *management succession* must be dealt with. Who will lead the merged firm and who will succeed that leader.[31] Third, *community concerns* about the merger should be anticipated and planned for. For example, the Bank Of America (BOA) and Security Pacific merger in 1992 ran into a hailstorm of community protests despite both banks having good CRA ratings. There were allegations that lending to low- and moderate-income neighborhoods would be severely curtailed by the combined bank, and that BOA would cut up to 25,000 jobs after the merger.[32] The merger was eventually approved, but only after considerable delay. Finally, information systems and corporate control issues are important. The merging partners should be careful to recognize that managerial incentive contracts that were effective prior to the merger may not be the best contracts to use after the merger.[33] Hence, internal incentive and control systems, as well as intrafirm communication channels, may need to be redesigned. These design issues should be dealt with prior to the consummation of the merger, and new systems should be put in place as soon after the merger as possible.

- **The Planning Process:** To have successful strategic plans, the bank must be willing to commit sufficient resources to the *planning process* itself. Citigroup and

31. For a fascinating account of how these concerns can sabotage merger plans, see Neuharth's (1989) discussion of how the proposed Gannett-CBS merger fell apart.

32. These allegations were denied by the banks concerned. BOA stated, prior to the Federal Reserve approval, that no more than 15,000 jobs would be lost.

33. See Ramakrishnan and Thakor (1991).

Wells Fargo may have led the way in legitimizing the planning process with the elevation of key senior executives to the position of executive vice president in charge of strategy and development. Strategic planning should not be an afterthought or a buzzword. It should be a way of defining what the bank is and what it wants to be.

- **Organizational Design:** There are two main issues in organizational design: the organization of the bank itself and the organization of the board of directors. Consider the *organization of the bank* first. In the better-managed banks, the trend is toward centralized corporate finance decision making. Indeed, finance is the single most centralized function in most corporations.[34] But centralized decision making does not mean that operational execution must also be centralized. It is often essential that information be gathered and decisions executed in a decentralized way to take advantage of insight into local markets. Once gathered, such information can then be made centrally available for optimal decision making by the central treasury. There are basically two approaches to the centralization of financial decisions: (i) assignment of worldwide business responsibility to the central finance function, which has direct access to the finance departments of all subsidiaries, and (ii) establishment of clear incentives for business units and subsidiaries to turn voluntarily for assistance to the central finance department, usually combined with the option of first refusal.

  In practice, the degree of centralization differs by product area. Capital market issues, mergers and acquisitions (M & A), and insurance usually are centralized, whereas short-term financing and payment transfers are often handled locally. *Figure 13.7* illustrates how one multinational bank divides its treasury functions. The contemporary model is flat organization, collegial governance, and functionally driven centralization/decentralization.

  We turn now to the *organization of the bank's board of directors*. The nature of the bank board is significantly different from that of a board for most industrial and service corporations.[35] In a general sense, a bank board becomes more involved in operations than an industrial board. Some specific differences are noted below.

  - A bank director can be fined up to $1 million per day by regulators for misbehavior. There is no comparable penalty for unregulated nonfinancial firms.
  - Unlike their counterparts in nonfinancial firms, bank directors have residence and citizenship requirements.
  - Investment bankers may not serve as bank directors; no comparable barrier exists in business corporations.
  - A bank director must make a financial commitment by taking an ownership position; business corporate law typically exacts no such requirements.
  - With exceptions, national bank directors may not serve on more than one national bank board; there is no comparable general prohibition in nonbanking corporations.

34. See Hagermann (1991).
35. See also Mueller (1990).

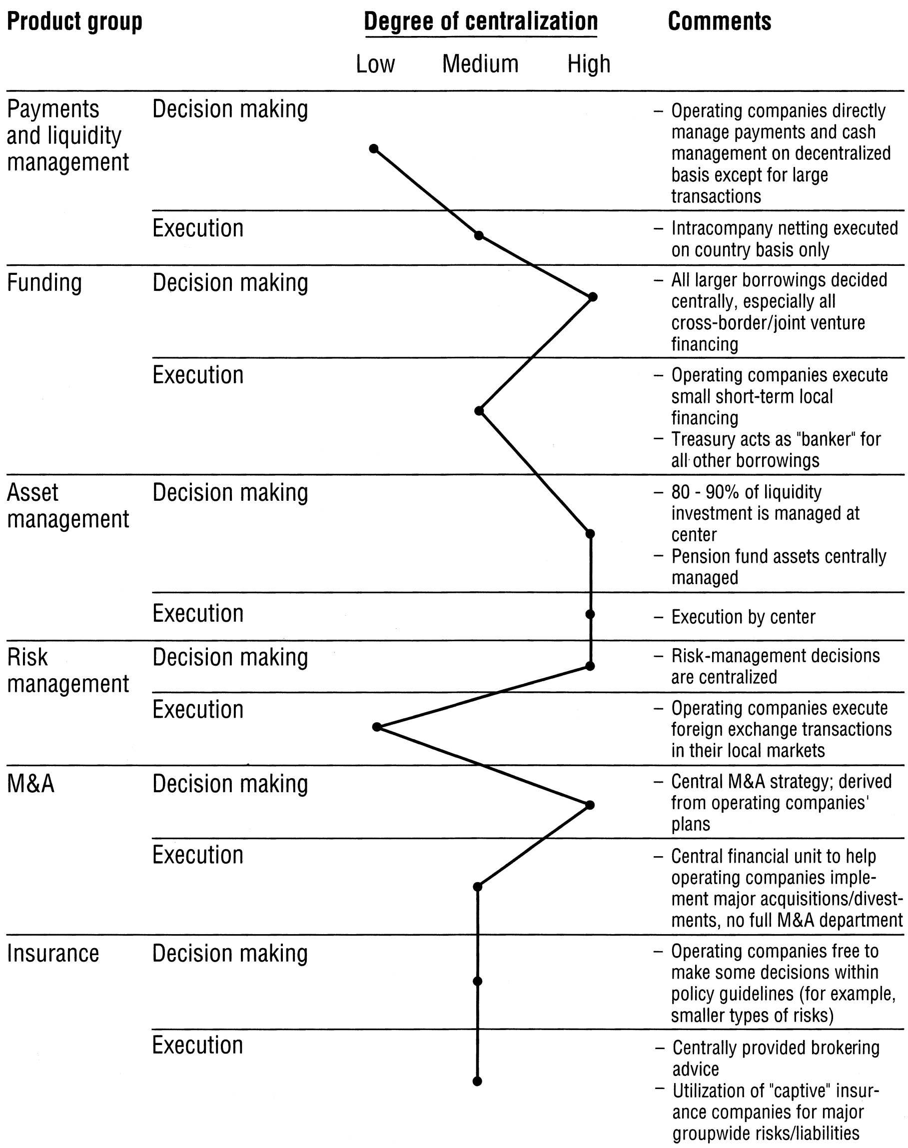


FIGURE 13.7 Degree of Centralization of Major Activities
Source: Hagermann (1991).

The formal organizational structure of a bank board is set forth in the statutes, in the corporate charter, and in the bylaws. There are two important features of any board. One is the relationship and interface between certain company officers, particularly the CEO and the treasurer. The other is the extent of each delegated power of the board. Corporate financial authority is normally described in the bylaws, and the accountability of each unit and its officers is defined. Directors perform specific roles by serving on committees. The five most common committees in U.S. bank boards are: (i) executive, (ii) salary and bonus, (iii) stock options, (iv) audit, and (v) finance.

Case Study

We now discuss two case studies to illustrate some of the concepts discussed thus far. One case relates to the management of risk and the other to the management of opportunities.

Risk Management Using RAROC and EVA

The Background: In the 1980s, Bankers Trust began to use a concept called "*Risk-Adjusted Return on Capital*" (RAROC), to employ a risk-management approach that mirrored sophisticated capital budgeting involving risk-adjusted discount rates. Basically, the required RAROC is the minimum expected return on the bank's equity capital that the bank should earn, given the risk in the investment it is considering. That is, it adjusts the required return on an investment for the amount of risk taken. It does so by measuring appropriate amount of capital to that activity—the greater the risk, the greater is the capital allocated, and the higher is the required rate of return. It is defined as:

$$\text{Required RAROC} = \frac{\text{Expected Net Income}}{\text{Risk-Adjusted Capital}}$$

$$= \frac{[\text{Spread Income} + \text{Fees}] - [\text{Overhead Expenses}] - [\text{Expected Losses}] - [\text{Taxes}]}{[\text{Amount of Equity Capital Allocated to the Loan}]}$$

We have already seen in Chapter 6 how many of these components are calculated. The expected loss on a loan = probability of default × the expected loss given default. The amount of equity capital allocated to the loan is an increasing function of two variables: the size of the loan and the loan-loss volatility (standard deviation of the loan loss).

To measure the actual value created for the bank's shareholders by a loan is then given by:

$$\text{Economic Value Added (EVA)}$$
$$= \left\{ \left[\frac{\text{Actual Net Income}}{\begin{matrix}\text{Amount of Equity}\\ \text{Capital Allocated}\end{matrix}} \right] - \text{Required RAROC} \right\} \times$$
$$\{\text{Amount of Equity Capital Allocated to the Loan}\},$$

$$\text{where Actual New Income} = [\text{Spread Income} + \text{Fees}] - [\text{Overhead Expenses}] - [\text{Actual Loan Losses}] - [\text{Taxes}]$$

The idea is to price each investment (Spread income + fees) so that it generates a positive EVA.

Implementation: RAROC and EVA work at two levels at banks that use them. First, they provide a reference point for individual employees. For example, a swap trader offered a deal will compare the return with the required RAROC. If the return falls below the required RAROC, the deal will either be repriced or turned down. The bank enforces strict limits on trading positions. For example, at Bankers Trust, in some markets, if a trader lost 10 percent of his RAROC capital in a month, he had to stop trading until the following month: a loss of 30 percent over a year led to a complete halt.

Second, RAROC is used at a portfolio level. Consider the swap dealer again. If a proposed deal reduces the overall riskiness of his trading book, he will need less capital and can thus find the deal profitable for the bank even with a lower return. RAROC thus allows for the benefits of diversification.

Many banks apply RAROC and EVA to loans as well as to securities. For Bankers Trust, in the 1980s, RAROC revealed that mainstream corporate lending was declining in risk-adjusted profitability. Competition precluded raising loan prices, so Bankers Trust reduced its loan book by nearly half between 1986 and 1992. RAROC also led to the bank deciding to sell many of its loans, so that loan origination (underwriting) profits were sustained without sacrificing capital.

The Spreading Faith: The RAROC concept has been adopted by many other banks. Examples are Citigroup, Chase Manhattan, Corestates Financial, Deutsche Bank, and Bank of America.

Opportunity Management: Bank of Hawaii

The Background: Bank of Hawaii, with $10 billion in assets in 2005, finished as the top performer in the Bank Performance Scorecard prepared by *Bank Director* with assistance from the investment banking firm of Sandier, O'Neill and Partners.[36] The bank had a Tier-1 capital ratio of 10.25 percent, a leverage ratio of 7.42 percent and a ratio of nonperforming loans to total loans and other real estate owned of merely 0.18 percent. How did a bank that in the late 1990s was performing so poorly that it was forced to operate under a memorandum of understanding with the Federal Reserve Bank of San Francisco and the FDIC manage to do so well?

The simple answer is that in 2000, the bank hired a new CEO, Michael E. O'Neill (former CFO, Bank of America Corporation), who decided that the bank would seek its growth opportunities *not* through the pursuit of commercial lending opportunities in diverse geographies in California and Asia, but by mining the rich Hawaiian economy for all it is worth. Thus, in a 3-year turnaround plan, the bank retrenched from California and all its Asian outposts and refocused on Hawaii, while also cleaning up its loan portfolio. The bank reduced its portfolio of syndicated loans, significantly tightened its underwriting standards, and built a strong balance sheet with capital ratios.

Going forward, the bank has plans to exploit numerous growth options. In addition to its sizeable retail and commercial banking operations, the bank also has several subsidiaries that are focused on investment management and equipment leasing, along with insurance and insurance agency services. The bank plans to tie these businesses together by cross-selling more effectively.

36. See Milligan (2005) for details.

Conclusion

In this chapter we have discussed the key issues in the management of a bank. We outlined a systematic framework for thinking about a bank's management challenges, and explained how these challenges can be ultimately linked to the basic economic functions served by a bank. Numerous concepts covered in previous chapters were highlighted to provide an integrated framework.

Review Questions

1. Explain the correlation between profit opportunities and risks in banking. What implications does this have for bank regulation?
2. Provide a taxonomy of the management of risks and opportunities in banks.
3. Which is the single most important risk for banks to manage and why? What sort of steps should banks take to ensure the efficient management of this risk?
4. Describe the evolution of risk-management techniques and how this evolution was affected by the changing nature of banking.
5. What is regulatory risk in banking? How does this risk affect the bank's cost of (equity) capital? What does regulatory risk imply about differences in the international competitiveness of banks in different countries?
6. What are the different types of EFT risks a bank faces? Have these risks become greater in recent years? (Be careful to distinguish between and relate idiosyncratic and systemic risks).
7. What are the different approaches to controlling EFT risks?
8. What is "netting" and how does it reduce operational risks in payment systems?
9. What are "credit swaps"? What are some of the impediments to the success of credit swaps? What are the similarities and differences between credit swaps and securitization?
10. What is the similarity between the risk of fraud and asset-substitution moral hazard?
11. What risks should the bank try to manage "in house" and what risks should it try to have externally insured? Explain your answer.
12. Analyze the process of crisis management in a bank.
13. Provide an in-depth discussion of strategic planning in a bank.
14. How is the role of the board of directors different in a bank from that in a nonfinancial firm? What constituencies is a bank board responsible to? If you could design a bank board, what kind of people would you put on it?

References

Barefoot, Jo Ann, Anjan V. Thakor, and Jess C. Beltz, "Common Ground: Increasing Consumer Benefits and Reducing Regulatory Costs in Banking," monograph published by *The Prochnow Educational Foundation*, 1993.

Bergman, Becky, "Redefining Customer Service: Where Customers Are King," *Bank Director*, 4th Quarter 2005, 34–41.

Besanko, David, and Anjan V. Thakor, "Relationship Banking, Deposit Insurance, and Bank Portfolio Choices," *Capital Markets and Financial Intermediation* (Xavier Vives and Colin Mayer, eds.), Cambridge University Press, 1993, 292–318.

Boot, Arnoud W.A., and Anjan V. Thakor, "Banking Scope and Financial Innovation," *Review of Financial Studies* 10-4, Winter 1997, 1099–1131.

——, "Moral Hazard and Ambiguity," Indiana University working paper, revised, June 1992.

——, "Can Relationship Banking Survive Competition?" *Journal of Finance*, April 2000, 679–713.

Boyd, John, and Mark Gertler, "U.S. Commercial Banking: Trends, Cycles, and Policy," NBER working paper, February 1993.

Carroll, Evelyn F., John P. Danforth, Carter H. Golembe, and P. Michael Laub, *The Burden of Bank Regulation*, a study prepared for *The American Banker's Association*, August 1989.

Chaney, Paul, and Kirk Phillipich, "Shredded Reputation: The Lost of Audit Failure," *Journal of Accounting Research*, 40-4, 2002, 1221–1245.

Consumer Access to Basic Financial Services, Hearing Before the Subcommittee on Consumer Affairs and Coinage of the Committee on Banking, Finance and Urban Affairs, House of Representatives, One Hundred First Congress, October 17, 1989.

Eisenbeis, Robert A., and Myron L. Kwast, "Are Real Estate Specializing Depositories Viable? Evidence From Commercial Banks," *Journal of Financial Services Research 5*, 1991, 5–24.

Engen, John R., "Many Happy Returns," *Bank Director*, 4th Quarter 2005, 26–28.

Greenbaum, Stuart I., and Bryon Higgins, "Financial Innovation," in *Financial Services: The Changing Institutions and Government Policy* (George Benston, ed.), Prentice Hall, Inc., Englewood Cliffs, N.J. 1983.

Hagermann, Helmut, "Should Your Company Be a 'Corporate Bank'?" *The McKinsey Quarterly* 4, 1991, 19–34.

Hellauer, Brian, "TCF Financial Finds Fees, Profits Catering to Mass Market," *American Banker's Management Strategies*, May 10, 1993.

Holland, Kelley, "Capital: C & S Takes Wary Approach to Rate Risk," *American Banker*, March 22, 1990.

Humphrey, David B., "Market Responses to Pricing Fedwire Daylight Overdrafts," *Economic Review*, Federal Reserve Bank of Richmond, May/June 1989, 23–34.

"International Banking: A Comedy of Errors," *The Economist*, April 10, 1993.

McCoy, Charles, "Credibility Gap: First Interstate Bankcorp Sheds Its Chairman But Not Its Problems," *Wall Street Journal*, January 18, 1990.

Mengle, David L., "Behind the Money Market: Clearing and Settling Money Market Instruments," *Economic Review* 78/5, Federal Reserve Bank of Richmond, September/October 1992, 3–11.

Mengle, David L., David B. Humphrey, and Bruce J. Summers, "Intraday Credit Risk, Value, and Pricing," *Economic Review* 73, Federal Reserve Bank of Richmond, January/February 1987, 3–14.

Milligan, Jack, "Banking in Paradise," *Bank Director*, 4th Quarter 2005, 31–32.

——, "What Makes a Good Bank Good?" *Bank Director*, 4th Quarter 2005, 16–24.

Mitroff, I., P. Shrivastava, and F. Udwadia, "Effective Crisis Management," *Academy of Management Executive*, 1987, 283–292.

Mueller, Robert K., "Roles and Responsibilities of the Board," Chapter 3 in *The Bank Director's Handbook*, Second Edition, 1990.

Neuharth, Allen, *Confessions of an S.O.B.*, Bantam Doubleday Dell Publishing Group, Inc., New York, 1989.

Parker, George, "Now Management Will Make or Break the Bank," *Harvard Business Review*, November–December 1981, 140–148.

Parkinson, Patrick, Adam Gilbert, Emily Gollob, Lauren Hargraves, Richard Mead, Jeff Stehm, and Mary Ann Taylor, "Clearance and Settlement in U.S. Securities Markets," Board of Governors of the Federal Reserve System, Staff Study 163, 1992.

Pierce, James L., and Samuel B. Chase, *The Management of Risk in Banking*, monograph, *Association of Reserve City Bankers*, July 1988.

Rajan, Raghuram, "Why Bank Credit Policies Fluctuate: A Theory and Some Evidence," *Quarterly Journal of Economics* 109, 1994, 399–442.

Ramakrishnan, Ram T.S., and Anjan V. Thakor, "Cooperation Versus Competition in Agency," *Journal of Law, Economics and Organization* 7–2, 1991, 248–283.

Reilly, Anne H., "Crisis Management in the Banking Industry: Strategic Issues for Managers," BRC working paper #170, Northwestern University, November 1988.

Sease, Douglas R., and Robert Guenther, "Financial Loss: Big Banks Are Plagued by a Gradual Erosion of Key Profit Centers," *Wall Street Journal*, August 1, 1990.

Song, Fenghua, and Anjan V. Thakor, "Relationship Banking, Fragility and the Asset-Liability Matching Problem," *Review of Financial Studies*, forthcoming.

Thakor, Anjan V., "Do Loan Commitments Cause Overlending?" *Journal of Money, Credit and Banking* 37–6, December 2005, 1067–1100.

PART ♦ VIII

Corporate Control and Governance

This page intentionally left blank

CHAPTER ♦ 14

Mergers and Acquisitions

"In the takeover business, if you want a friend, you buy a dog."
Carl Icahn

Glossary of Terms

TBTF: Too Big to Fail (see Chapter 10).

BHCA: Bank Holding Company Act.

BMA: Bank Merger Act.

FIRREA: Financial Institutions and Regulatory Reform Enforcement Act (see Chapter 11).

OCC: Office of the Comptroller of the Currency (see Chapters 11 and 12).

FDICIA: Federal Deposit Insurance Corporation Improvement Act (see Chapters 11 and 12).

Intramarket Merger: A merger of two firms operating in the same market, competing for the same customers.

Intermarket Merger: A merger of two firms operating in different markets.

Horizontal Integration: When firms providing similar services or manufacturing similar products merge. This is an expansion of scale.

Vertical Integration: When a firm merges with a supplier or a customer. This is an expansion of scope.

Zero Abnormal Returns: When the change in a firm's stock price yields a return that is consistent with the firm's systematic risk, that is, the return predicted by the Capital Asset Pricing Model.

Introduction

The merger trend, observed across the entire cross-section of industries in the U.S. economy and elsewhere during 1980–2000, has been particularly evident in the financial services industry. Since 1980, the U.S. banking industry has experienced a sustained and unprecedented level of merger activity that has substantially affected banking structure. From 1980 through 1998, there were approximately 8,000 mergers, involving about $2.4 trillion in acquired assets. Although there were more bank mergers during the 1980s than during later periods, the period of the 1990s, particularly 1995–98, was characterized by an increasing number of very large mergers, including some that surpassed in size any previous mergers. Even though the pace of U.S. banking merger activity slowed down during 2001–2003, it picked up again in 2004, which turned out to be the highest since 2000. There were 271 deals in 2004, totaling $131.5 billion in transaction value. Consolidations are wringing out excess capacity and reducing the number of firms. At almost 7,500 banks in 2006, the U.S. is still considered by some to have too many banks, and the end of banking consolidation has probably not yet been witnessed.

In other countries, there are striking cost inefficiencies that remain to be eliminated. Banks in major European countries can still lower their cost/income ratios, and corporate governance issues and shareholder-wealth maximization are gaining prominence in Japan. A study of mergers and acquisitions in banking is, therefore, both timely and important. In this section we will discuss a variety of issues related to corporate control contests in banking.

The rest of the chapter is organized as follows. In the next section we examine recent trends in mergers and hostile takeovers in banking. After that we examine the literature on corporate control in general, and follow this with an examination of its implications for bank mergers. The next section takes up hostile takeovers in banking.

Recent Trends in Mergers and Acquisitions in Banking

Prior to 1985, bank mergers in the U.S. occurred mainly on an *intrastate* basis, and merger activity was heaviest in states that had recently liberalized their unit banking laws or bank holding company restrictions. Many of the transactions were small and involved stock deals. However, once the Supreme Court affirmed regional *interstate* compacts on June 10, 1985, the floodgates were opened. Most states permitted some form of interstate banking after that and many permitted nationwide banking. The passage of the Interstate Banking Act of 1994 provided further impetus. The largest deals are now interstate combinations. Merger activity in banking reached its high-water mark for the decade of the 1980s in 1986, declining thereafter, and picking up again in 1989. The 1990s opened with a sharp drop in mergers, but since 1994 merger activity in banking has been very high. Figure 14.1 plots this activity for 2000–2005, and Table 14.1 lists the major mergers in banking announced in 2004.

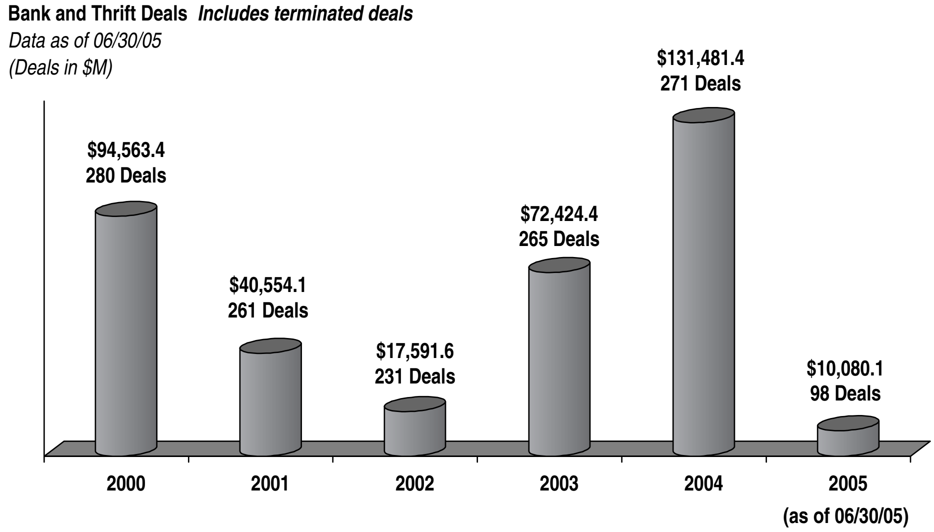


FIGURE 14.1 Bank and Thrift Deals 2000–2005 (Including Terminated Deals)
Source: Charlottesville, VA-based SNL Financial, snl.com and *Bank Director*, 3rd Quarter, 2005.

Table 14.2 lists the top financial advisers in bank and thrift deals in 2004 and Table 14.3 lists the top legal advisers.

Corporate Control Issues

Reasons to Merge: The Whole Exceeds the Sum of the Parts

Two firms may merge because the value of the combined entity exceeds the sum of their stand-alone values. Such value enhancement can arise from a variety of sources.

One such source is *economies of scale*, which means that average cost declines with volume of output. A second source is *economies of scope*, which means that average cost declines as the scope of activities increases. A third source is *elimination of redundant capacity*, which means that banks are able to eliminate waste by merging. A fourth source is *increased market power* achieved through horizontal or vertical integration. Vertical integration provides increased power in the factor input markets and/or greater control over distribution channels, resulting in lower costs. Horizontal integration provides increased market power with respect to customers, resulting in higher profits. A fifth source is *improvement in managerial efficiency*. The idea is that a merger can help to remove inefficient managers and replace them with efficient managers. Moreover, it can also permit better incentive contracts to be written for managers. Sixth, mergers can enhance value by providing an *opportunity to enter new markets, cross-sell products, and diversify*. Finally, banks may merger even when there

TABLE 14.1 Top 10 Bank & Thrift Deals Announced in 2004 (Ranked by Deal Value at Announcement) *Data as of 05/25/05*

Buyer	**City**	**Target**	**City**	**Announce Date**	**Completion Date**	**Deal Value ($M)**	**Price/ Book (%)**	**Price/Tangible Book (%)**	**Price/LTM Earnings (x)**
J.P. Morgan & Chase	New York, NY	Bank One Corporation	Chicago, IL	01/14/2004	07/01/2004	58,783.3	256.10	285.43	17.29
Wachovia Corporation	Charlotte, NC	South Trust Corporation	Birmingham, AL	06/21/2004	11/01/2004	14,365.4	306.67	372.82	19.64
Royal Bank of Scotland Group, Plc	Edinburgh	Charter One Financial, Inc.	Cleveland, OH	05/04/2004	08/31/2004	10,526.9	305.63	354.30	19.18
Sun Trust Banks, Inc.	Atlanta, GA	National Commerce Financial Corp.	Memphis, TN	05/07/2004	10/01/2004	7,432.5	249.92	451.37	20.79
North Fork Bancorporation, Inc.	Melville, NY	GreenPoint Financial Corp.	New York, NY	02/15/2004	10/01/2004	6,396.3	297.53	370.80	12.14
Regions Financial Corporation	Birmingham, AL	Union Planters Corporation	Memphis, TN	01/22/2004	07/01/2004	6,000.8	NM	NM	NM
Capital One Financial Corporation	McLean, VA	Hibernia Corporation	New Orleans, LA	03/07/2005	NA	5,351.3	261.94	323.25	17.74
TD Bank Financial Group	Toronto	Banknorth Group, Inc.	Portland, ME	08/25/2004	03/01/2005	3,818.1	234.53	477.69	17.80
National City Corporation	Cleveland, OH	Provident Financial Group, Inc.	Cincinnati, OH	02/16/2004	07/01/2004	2,133.9	220.82	245.82	21.14
Fifth Third Bancorp	Cincinnati, OH	First National Bankshares of Florida, Inc.	Naples, FL	08/02/2004	01/01/2005	1,530.2	263.07	NA	42.13

Source: Charlottesville, VA-Based NSL Financial, SNL.com and *Bank Director*, 3rd Quarter, 2005. See Milligan (2005).

TABLE 14.2 Top Financial Advisors in Bank & Thrift Deals in 2004 (Ranked in Total Deal Value) *Data as of 1/1/04–6/30/05*

Rank	Firm	Number of Deals	Total Deal Value ($M)
1	Lazard Freres & Co.	1	58,783.3
2	UBS Securities LLC	6	30,811.8
3	Merrill Lynch & Co.	4	28,365.4
4	Lehman Brothers Inc.	12	25,295.5
5	Goldman Sachs & Co.	6	24,320.2
6	J.P. Morgan Securities Inc.	5	19,502.9
7	Sandler O'Neill & Partners LP*	52	15,546.4
8	Keefe Bruyette & Woods, Inc.	58	14,867.8
9	Morgan Stanley	3	8,556.5
10	Credit Suisse First Boston USA	3	6,145.3
11	Citigroup Global Markets	3	3,625.7
12	Hovde Financial LLC*	28	3,610.8
13	Ryan Beck & Co.	14	3,383.5
14	Secura Group	1	2,133.9
15	SunTrust Robinson Humphrey	4	2,046.2
16	Banc of America Securities	1	850.0
17	McDonald Investments Inc.	7	759.3
18	Northeast Capital & Advisory	2	758.3
19	Castle Creek Financial	7	780.4
20	Friedman Billings Ramsey & Co	9	716.5

Source: Charlottesville, VA-Based SNL, Financial, snl.com and *Bank Director*, 3rd Quarter, 2005
*Deal Value for at least one deal is NA.

are *no* gains for their shareholders, simply because it generates private benefits for the CEOs of the merging firms.[1]

Mergers Come About in Two Ways: Friendly Mergers and Hostile Takeovers

Firms can combine their operations either through a friendly merger in which those in control of the two firms come to a mutual agreement, or through a hostile takeover in which the acquiring firm gains control without the cooperation of the target's managers. We will refer to friendly mergers as just "mergers" and to hostile takeovers as just "takeovers" for short. Mergers and takeovers focus on different sources of gains in combining the two firms and have different implications for stock price reactions to the announcement of a combination.

1. Boot, Milbourn and Thakor (1999) use this framework to explain megamergers that expand bank size as well as scope.

TABLE 14.3 Top Legal Advisors in Bank & Thrift Deals in 2004 (Ranked by Total Deal Value) *Data as of 1/1/04 to 6/30/05*

Rank	Firm	Number of Deals	Total Deal Value ($M)
1	Wachtell Upton Rosen & Katz	14	101,577.8
2	Simpson Thatcher & Bartlett	2	62,601.4
3	Sullivan & Cromwell LLP	14	23,323.1
4	Bradley Arant Rose & White	4	14,599.6
5	Burr & Forman	1	14,365.4
6	Skadden Arps*	4	14,241.6
7	Goodwin Procter LLP*	6	11,948.7
8	Silver Freedman & Taff LLP*	8	11,322.2
9	Davis Polk & Wardwell	2	10,738.0
10	Alston & Bird LLP	7	7,788.8
11	Bass Berry & Sims Plc.	3	7,445.8
12	King & Spalding	1	7,432.5
13	Elias Matz Tiernan & Herrick	11	4,957.1
14	Osier Hosking & Harcourt	1	3,818.1
15	Stevens & Lee PC	4	2,479.8
16	Arnold & Porter	1	2,133.9
16	Jones Day	1	2,133.9
17	Smith Gambrell & Russell	5	1,868.8
18	Pillsbury Winthrop Shaw	4	1,779.5
19	Graydon Head & Ritchey LLP	1	1,530.2

Source: Charlottesville, VA-Based SNL, Financial, snl.com and Bank Director, 3rd Quarter, 2005
*Deal Value for at least one deal is NA.

Separation of Ownership and Control and Takeovers

When a firm makes an offer to acquire another firm, the offering price is invariably greater than the market price of the target firm prior to the offer. Why then would the target ever resist the takeover attempt? Separation of ownership and control in publicly traded firms provides much of the answer.[2] These firms are owned by their shareholders but operated by managers. Shareholders are therefore principals who delegate (via boards of directors) to their agent, the manager, the task of operating the firm.[3] Since the manager will operate the firm to maximize his or her own expected utility, an *agency problem* arises (recall Chapter 1). This agency problem manifests itself in various ways, which are discussed below.

- **Excessive Conservatism in Investment Decisions:** Managers may choose less risky projects than is optimal for the shareholders. This can sacrifice firm value. Such conservatism may arise from managerial risk aversion (the manager is concerned about losing his or her job), or from managerial concerns about how

2. In addition to this, there may be strategic reasons for targets to resist takeovers initially to negotiate for a better price.

3. See Jensen and Meckling (1976) for a discussion of the consequences of this separation of ownership and control. Mester (1989) provides a discussion of the agency problem in banking.

firm performance may impinge on the manager's rewards, including reputation and career concerns.[4]

- **Excessive Consumption of Perquisites:** This wastes resources and diminishes firm value. Moreover, it sets a poor example for subordinates.
- **Pursuit of Negative NPV Projects:** This can take many forms and occur for different reasons. Managers may choose new negative NPV projects because these projects produce private, noncontractible managerial benefits, or simply because there is "free cash flow"—cash left over after all positive NPV projects have been exhausted—that can be used for "empire building."[5] That is, the manager may gain from an increase in size because his or her compensation is tied to size. Alternatively, managers may hang on to money-losing projects even though liquidating/selling them would be best for shareholders. This can happen because managers wish to protect their reputations by delaying disclosure to the market that they erred in having chosen these projects.[6]
- **Managerial Entrenchment:** The manager may know very well that he or she is inefficient and that there are others—some possibly within the firm—who would do a better job of maximizing shareholder wealth. The manager has an obvious interest in blocking potential successors within the firm and keeping at bay those outside the firm. Thus, resistance to takeover attempts can be viewed as a manifestation of the agency problem between shareholders and managers.

Mechanisms for Reducing the Managerial Agency Problem

We briefly discuss the three ways shareholders can align managerial incentives with their own.

- **Managerial Compensation Schemes:** These can be designed to provide a stronger value-maximization motive for the manager. For instance, the manager can be compensated with stock and stock options that have trading restrictions.[7] Considerable research on the relationship between firm value and managerial ownership of the firm finds that higher managerial ownership of the firm leads to better decisions for shareholders.[8]
- **Monitoring:** Shareholders monitor managers and can replace them if they appear untalented or unmotivated. However, monitoring is usually delegated by shareholders to the board of directors. The effectiveness of the board is hindered by informational asymmetries between the chief executive officer (CEO) and the board as well as the presence of "inside" directors who may

4. See Hirshleifer and Thakor (1992) for a reputational analysis of managerial conservatism in investment decisions, even when managers are risk neutral. See also Milbourn, Shockley and Thakor (2001). The seminal paper on this is Holmstrom (1999).

5. See Jensen (1986) for an examination of the consequences of free cash flow.

6. See Boot (1992) for a model along these lines.

7. Trading restrictions ensure that the equity will be retained by the manager in his or her portfolio.

8. See, for example, Cotter and Zenner (1994), and Martin and McConnell (1989).

either be senior executives of the firm or members appointed by the CEO.[9] This creates a reason for institutional shareholders and large creditors, like banks, to monitor management.

- **Takeovers:** If the manager is wasting the firm's resources, its market price should reflect this inefficiency. Indeed, one can think of the market value of the firm, V_m, as consisting of two parts: (i) the value of the firm under current management (V_c), multiplied by the probability that the firm remains independent, and (ii) the value of the firm if it is taken over (V_t), multiplied by the probability of that event occurring.[10] Inefficiency in managing the firm can reduce its value and result in a lower market price, holding fixed the probability that current management will remain in control. Another firm can then offer a price higher than the current market price and still expect to earn a profit for its shareholders as long as the offer price is between the market price and V_t. Thus, a takeover can be profitable for the acquiring firm and benefit the target shareholders as well, even without technological or market power synergies from the merger.

Factors Limiting the Effectiveness of Takeovers in Disciplining Management

These three corporate control mechanisms operate jointly. The more effective are compensation schemes and board monitoring, the less likely is a takeover. However, even when there is a role for takeovers, their effectiveness can be impeded by a *free-rider* problem.[11] Imagine a potential acquirer that bids a price V_b for the target, with $V_m < V_b < V_t$. In order for the bid to be successful, a sufficient number of target shareholders must offer to sell their shares to the bidder. Suppose for the moment that all target shareholders sell their shares, so that the takeover is successful. Can this be a Nash equilibrium among noncolluding shareholders? To answer this, consider the strategy of an individual target shareholder contemplating whether to sell his or her shares. If the shareholder owns a percent of the firm, then the shareholder receives αV_b if he or she sells. But if the shareholder retains his or her shares when all others sell, then the value of his or her holdings is αV_t. Since $V_t > V_b$, this shareholder is better off holding on to his or her shares, conditional on all the other shareholders selling. But this is true for every shareholder! Thus, nobody sells and the takeover attempt fails as each shareholder has an incentive to "free ride" on the willingness of the other shareholders to sell. One way for the bidding firm to induce all target shareholders to tender their shares is to bid V_t, in which case there is no profit from the takeover. So if the bidder has to expend some cost, C, at the outset to become informed about the target's value under new management, it will be unable to recoup this cost. It will consequently not make this initial investment in information acquisition, and there will be no bids made for the target firm.

9. Hirshleifer and Thakor (1993, 1994) theoretically examine the role of the board of directors in disciplining management. Song and Thakor (2006) examine the interaction between the career concerns of the CEO and the board.

10. Stulz (1988) develops a model along these lines.

11. This issue was explained by Grossman and Hart (1980).

One way to avoid this free-rider problem is for a potential bidder to secretly acquire a sufficiently large fraction of the target firm prior to announcing its intention to purchase control.[12] Suppose the potential bidder discretely acquires a fraction β of the outstanding shares. If the acquisition is carried out in complete secrecy, the potential bidder will pay βV_m for these shares. After this ownership fraction is reached, suppose a bid is made for the remaining shares for a total amount $(1 - \beta)V_t$. Then the target shareholders have no reason not to tender their shares because their payoff can never be higher from holding on to their shares. And even though the bidder earns no profit on the remaining shares purchased to acquire the target, it realizes a profit on the shares acquired prior to making public its intent to take over the target. This profit is $\beta[V_t - V_m]$. As long as $\beta[V_t - V_m] > C$ (the cost of effecting the transaction, including the information acquisition cost), the bidder's incentive to acquire information about the target is sustained.

Although the free-rider problem can be resolved, the effectiveness of takeovers in disciplining target management can be reduced by takeover resistance. For example, managers in a target can "swallow a *poison pill*" that makes the target unattractive to potential bidders. For example, the target may acquire another firm in order to increase the likelihood of antitrust litigation if its potential acquirer succeeds. Other poison pills are financial restructuring that raises the cost of a takeover, or selling off some assets that attracted the bidder.[13] The empirical evidence is that share prices of target firms decline on announcements of defensive restructurings in response to takeovers and on announcement of poison-pill antitakeover measures.[14] The conclusion is that takeovers result in gains to shareholders, but self-serving managers may impede the realization of these gains.[15] Mergers and takeovers are capital-market mechanisms for reconfiguring portfolios of real and financial assets and putting these assets to their highest-value uses.

Price Reactions to Mergers and Takeovers: The Theory

Management efficiency can affect both the price at which the firm's stock is traded and the price reaction to a takeover bid by another firm. We now analyze the manner in which this effect is manifested.

A merger/takeover has two possible objectives. One is to realize synergies between the merging partners. These potential gains are independent of target firm management. Let $S > 0$ represent synergy gains. A second important objective is to remove

12. This is based on Shleifer and Vishny (1986). Disclosure rules of the Securities and Exchange Commission (SEC) require that the reason for acquiring shares must be made public when 5 percent ownership is reached.

13. Of course, even though poison pills enable managers to sometimes hold on to their jobs, the threat of takeovers may result in actions that benefit shareholders. Boot (1992) argues that takeover threats can be effective in inducing managers to divest money-losing divisions they would have otherwise held on to. Jensen (1986) suggests that takeover threats induce managers to pay out free cash flow as dividends rather than investing in negative NPV projects.

14. See Jensen and Warner (1988) for an extensive discussion of this empirical evidence.

15. The Federal Reserve has taken the position that it will treat takeover bids the same as merger bids in assessing whether to permit a takeover. Alan Greenspan, then chairman of the Federal Reserve Board, made the following statement in testifying before the Senate Banking Committee in February 1988 on Bank of New York's takeover bid for Irving Bank, "Let's remember when we talk about hostile takeovers, the hostility is between the managements of the two organizations, not between the shareholders of either. In fact, the problem that exists is that too often, in my judgment, the managements try to protect themselves from, in effect, their own shareholders, who are essentially their bosses." See Mester (1989).

inefficient management in the target. Let $M > 0$ be the gain from removing inefficient management in the target firm.

Consider now two firms that are identical in all respects except management efficiency. One firm has inefficient management; we will superscript variables related to this firm by I. The other firm has efficient management; we will superscript variables related to this firm by E. Let V^I represent the true value of the target firm's assets under inefficient management. Then V^E, the value of the same firm's assets under efficient management, is given by $V_E = V_I + M$.

If there is a takeover, then *only* the inefficient target management will be replaced.[16] The true value of the target firm's assets after the merger (indicated by the subscript t) will reflect both the synergy gains as well as the gains from improved management. That is,

$$V_t^I = V^I + M + S \qquad [14.1]$$

$$V_t^E = V^E + S = V^I + M + S \qquad [14.2]$$

will be the postmerger asset values of the inefficient and efficient target firms, respectively.

Let q (where $0 \leq q \leq 1$) represent the probability that a potential acquirer will bid for the target firm and p (where $0 \leq p \leq 1$) the probability that the bid will succeed. Assume that the potential acquirer, in anticipation of the free-rider problem discussed earlier, is expected to bid V_t for the target. Then, assuming that the efficiency of each firm is common knowledge, the market values of the two firms, prior to the announcement of a takeover bid or merger, will be:

$$V_m^I = pq[V^I + M + S] + [1 - pq]V^I, \qquad [14.3]$$

$$V_m^E = pq[V^I + M + S] + [1 - pq][V^I + M]. \qquad [14.4]$$

That is, each market value is a weighted average of the pre- and post-merger asset values.

Now, if the management of the firm is efficient, it should display no aversion to merging and realizing the synergy gains S because it will not be replaced after the merger. That is, we would expect such firms to announce a merger, rather than a takeover attempt by the potential acquirer. So $p = 1$ for a merger announcement. On the other hand, if the target management is inefficient, it fears for its job when confronted with a merger possibility. Such a management will resist a takeover. Of course, takeover resistance may fail, but it will make the outcome of the takeover

16. The assumption that inefficient target management will be replaced after a takeover is empirically true. Management turnover in targets of takeovers is abnormally high. Martin and McConnell (1989) report a 55 percent turnover rate among CEOs, chairpersons, and presidents of target firms in successful takeovers, within the first year after the takeover.

attempt uncertain. Let us capture this uncertainty by assuming that $0 < p < 1$ for a takeover attempt. Given that it is common knowledge which management is efficient and which is not, we can write (14.4) as:

$$V_m^E = q[V^I + M + S] + [1 - q]V^I. \qquad [14.5]$$

Note that $V_m^E > V_m^I$, that is, the efficiently managed firm trades at a higher price.

Now suppose the uncertainty about whether a bidder will arrive is resolved and a potential acquirer bids for the target. The efficient firm's management will accept the bid and a merger will be announced. The inefficient firm's management will resist the bid, but the market price will move nonetheless due to resolution of the uncertainty about the bidder's arrival. The new market prices of the two firms (denoting these new prices with hats) will be:

$$\hat{V}_m^I = p[V^I + M + S] + [1 - p]S, \qquad [14.6]$$

$$\hat{V}_m^E = V^I + M + S. \qquad [14.7]$$

We now wish to compute the *relative price change* in response to the takeover/merger bid for each firm. Let RPC^I and RPC^E represent these relative price changes for the inefficient and efficient firms, respectively. Then,

$$RPC^I = [\hat{V}_m^I - V_m^I]/V_m^I \quad \text{and} \quad RPC^E = [\hat{V}_m^E - V_m^E]/V_m^E. \qquad [14.8]$$

That is,

$$RPC^I = \frac{p[1 - q][M + S]}{V^I + pq[M + S]}, \qquad [14.9]$$

$$RPC^E = \frac{[1 - q]S}{V^I + M + qS}. \qquad [14.10]$$

With a little algebra, we see that the relative price change will be greater with the announcement of a takeover than of a merger (that is, $RPC^I > RPCE$) if:

$$pM[M + S + V^I] > [1 - p]SV^I. \qquad [14.11]$$

Note that (14.11) will hold when M, the cost of management inefficiency, is sufficiently large and/or S, the synergy gains, are sufficiently small. This is intuitive. It says that if the value enhancement from replacing inefficient management in the target is sufficiently high, the relative price change associated with a takeover attempt will be greater than that associated with a proposed merger. Whether (14.11) holds in practice is an empirical issue. If the inequality in (14.11) is reversed, the price reaction to a merger is predicted to be greater than that to a takeover.

Mergers in Banking

To understand the increased pace of mergers and acquisitions in banking, we first discuss the changes in the legal and regulatory environment in which banks operate, and then analyze banks' motives to merge.

Legal and Regulatory Environment

The two major federal regulations affecting bank mergers and takeovers are the *Bank Holding Company Act* (BHCA of 1956) as amended in 1970, and the *Bank Merger Act* (BMA of 1960) as amended in 1966. We discuss each below.

The BHCA of 1956: The BHCA applies to bank holding companies (BHCs). A BHC controls one or more banks, where "control" is defined as ownership of 25 percent of any class of voting stock. The BHCA requires approval by the Board of Governors of the Federal Reserve before a bank can be merged with another bank or become a subsidiary of a BHC. Moreover, Regulation Y, which deals with changes in bank control, stipulates that any merger of a subsidiary of a bank holding company requires approval of the surviving bank's federal supervisory agency. Thus, if the surviving bank is a national bank, the OCC also will rule on the merger. Apart from the various bank regulatory agencies, the Justice Department has *guidelines* regarding the permissible market shares of the merging firms. The major purpose of the BHCA is to restrict the BHC to activities that neither have anticompetitive effects nor compromise the bank's safety.

In 1966, the BHCA was made consistent with the antitrust provisions of the BMA. In determining whether to approve a BHC's request for a new bank acquisition, the Fed was now directed to consider four criteria: (i) the convenience, needs and welfare of the community, (ii) the public interest, (iii) the soundness of the banks involved and the impact on the soundness of the industry, and (iv) the impact on competition. However, the BHCA applied only to multibank holding companies. Consequently, banks formed one-bank holding companies to engage in nonbanking activities. To eliminate this loophole, Congress amended the BHCA in 1970 to encompass one-bank holding companies as well. Another milestone in regulation was reached with the adoption of FIRREA in 1989. Prior to FIRREA, the acquisition of thrifts by BHCs was limited to special cases, such as failing thrifts. As part of FIRREA, Congress directed that BHCs be permitted to acquire thrift institutions (healthy or otherwise) without restrictions.

In addition to banking, BHCs have been permitted to engage in commercial finance, mortgage banking, consumer finance, securities brokerage, leasing, data processing, and insurance underwriting. Apart from the ability to engage in such nonbanking activities, BHCs also offer numerous tax advantages.[17] First, dividend payments from banks to BHCs are not taxable. Second, a BHC can file a consolidated return for the entire organization so that losses in one subsidiary can be used to offset profits earned by the bank, thereby reducing the total tax liability. And finally,

17. See Eisenbeis (1983), Hawke (1989), Huber (1993), and Lash (1987).

the BHC can use debt (with tax-deductible interest) to fund the purchase of stock issued by a subsidiary bank, thereby substituting debt for equity in raising capital.

The BMA of 1960 and the Riegel-Neal Interstate Banking Act of 1994: First passed in 1960, the BMA called for the review of all bank mergers by the appropriate regulatory authority. In 1963, however, the OCC approved the merger of Philadelphia National Bank and Girard Corn Exchange Bank and Trust Company on the grounds that it would improve the convenience and needs of the community. The Department of Justice challenged the approval, charging that it would have anticompetitive effects. The case went all the way to the Supreme Court where the merger was rejected. In 1966, Congress attempted to clarify matters by amending the BMA. The Department of Justice was no longer assigned an advisory role. Rather, it could independently challenge the merger, and this input—although not binding—had to be officially considered by bank regulators. Thus, the legal standards for bank and BHC mergers are identical to those for nonfinancial firms and are patterned on traditional antitrust standards, relying on the *Sherman and Clayton Acts*. Both acts prohibit mergers that threaten competition, which is interpreted to include *extant* as well as *potential* competition. The concept of potential competition applies to firms that do not compete in the same market, but might conceivably do so. The potential entrant presumably influences prices toward more competitive levels. Markets characterized by potential competition will exhibit "limit pricing," as opposed to monopolistic pricing. Limit pricing is defined as the minimum price needed to deter entrants. It was believed that the Clayton Act could be used to deny market extension acquisitions that might discourage potential competition.

The doctrine of potential competition was never of great importance in banking, but waned even further as an impediment to mergers in the 1980s. A large number of intrastate market extension acquisitions that might have been challenged under the potential competition rule were approved during the 1980s.

One reason why fewer bank mergers are being denied on competitive grounds that they might reduce competition is that bank regulators are permitted to make exceptions. Regulators are allowed to approve mergers whose anticompetitive effects are outweighed by the benefits of the transaction in meeting the convenience and needs of the community. Additionally, they are able to approve mergers in which the "failing firm" exception to the antitrust laws can be applied. For example, the Federal Reserve allowed Norwest, a BHC that was four times larger than its nearest rival in the Minneapolis-St. Paul market, to acquire a failed, local thrift. While recognizing the anticompetitive effect of the merger, the Federal Reserve said that this was outweighed by the public benefit of saving the FDIC the cost of closing a negative-net-worth institution.

Another major factor in the increased pace of bank corporate control contests was the desire of regulators and politicians to not prevent interstate banking. Between 1982 and 1989, for example, about 400 interstate acquisitions were processed by the Federal Reserve. 1991–92 witnessed some spectacular mergers, such as those between: Chemical Bank and Manufacturers Hanover Corporation, Bank of America and Security Pacific, and NCNB Corporation and C&S/Sovran Corporation. This activity has picked up considerably since the Riegel-Neal Interstate Banking Act of 1994 repealed the branching restriction of the *McFadden Act*. Of course, even though the legal and regulatory environment may permit mergers, banks themselves must find it beneficial to merge. What factors contribute to banks' desire to merge? We turn to this issue next.

Bank's Incentives to Merge

The incentives that banks have to merge are similar to those for nonfinancial firms. In addition, banks have benefited from mergers because of the "Too Big to Fail" (TBTF) doctrine; FDICIA has sought to make this issue moot (see Chapter 12). The discussion below is an elaboration of the reasons to merge discussed in the previous section on corporate control.

- **Economies of Scale, Economies of Scope and the Elimination of Redundant Capacity:** The most popular reasons for merging are scale economies, scope economies, and elimination of redundant capacity. The logic is appealing. In theory, an intramarket merger should be able to sustain the same amount of business and dispense with one headquarters, one personnel department, one trading room, and—if they compete in the same local market—many branches. Effective "back-office" consolidation can achieve cost efficiencies as well. Many banks are entrusting this task to a new breed of techno-bankers, called "consolidators," who specialize in squeezing operating costs out of acquired banks. Their success or failure can dramatically affect a bank's stock price and its status as an acquiring institution.[18] Back-office consolidation involves streamlining information and communication systems and paring staff. Multiple data centers and software systems are consolidated into integrated units to conduct profitability analysis, manage demand-deposit transactions information, and keep track of retail banking business.

  Although savings due to back-office consolidation usually account for only a fraction of the total expected expense cuts,[19] most bank executives feel that back-office consolidation poses the greatest challenge in terms of management, implementation, and cost control. It is not surprising, therefore, that when a merger fails to achieve the promised cost savings, one of the contributing factors is usually a lack of effective back-office consolidation.

  Whereas intramarket mergers offer savings via consolidation and reconfiguration of backroom, support, and branch systems, intermarket mergers provide potential scope as well as scale economies.

  However, academic research has uncovered little evidence of either scope or scale economies in banking. Moreover, mergers do not seem to reduce costs or cut excess lending capacity. Although the evidence on post-merger operating performance is a bit mixed, most studies find that mergers do *not* lead to significant changes in performance.[20] The principal findings are as follows:

  - The average cost curve in banking is relatively flat.[21]
  - Merging banks experienced small but significant decreases in costs during the third and fourth postmerger years. However, this decline can be explained by an overall industry decline in expense ratios. That is, trends in noninterest

18. See Crockett (1992). To quote Mr. Ladd Willis, managing vice president of First Manhattan Consulting Group in New York, "You can only acquire if you can deliver on the expected operational cost savings. What these people do is absolutely critical."

19. For example, in the case of the merger of Chemical with "Manny Hanny," total projected savings after the merger were: $200 million from back-office consolidation, $100 million from cutting back on branches and real estate owned, and $350 million from staff reductions. See Kantrow (1991) for details.

20. See Rhoades (1994, 2000).

21. See Humphrey (1990).

expenses for merged banks have not been significantly different from those for nonmerged banks.[22] The only exception seems to be the largest mergers; cost efficiencies have been realized in these.[23]
  - On average, merging banks have not achieved significant economies in consolidating back-office operations.[24] However, there are mergers that result in cost savings for banks, and in these cases, loan spreads narrow after the merger. Moreover, the greater the cost savings, the stronger is the effect on the narrowing of loan spreads.[25]
  - No appreciable cost savings have accrued from multiproduct production by banks. That is, no significant scope economies appear to exist.[26]
  - Mergers do not reduce excess lending capacity. Banks that merged between 1980 and 1990 did not slow their growth significantly. The only effective means of reducing capacity seems to be via failure.[27]
  - Executive compensation increases after a merger and more so than if the growth occurs organically.[28]

- **Market Power:** Banks often merge to increase their market power. Many believe that this is the most important factor in intramarket mergers. This hypothesis is supported by evidence of mergers in the 1970s that credit availability, loan losses, deposit service charges, and interest-rate risk rose for banks after mergers.[29] Moreover, bank mergers produced neither significant service benefits nor significant reductions in service costs to the public.

- **Improvement in Managerial Efficiency:** A possibly significant source of gains in mergers is improved management efficiency. This improvement can come in three forms. First, with takeovers, inefficient managers in the target banks are likely to be replaced by the acquiring banks. Second, the very threat of a takeover can improve management efficiency. And third, mergers often lead to diversification of the asset bases and funding sources of the merging banks. This diversification means that there is less "noise" in the combined entity's performance measures, including the stock price. Hence, the performance of the bank is more closely tied to the efforts of its managers.[30] Empirical evidence indicates that very large bank mergers result in significant efficiency gains.[31]

- **Too Big To Fail (TBTF):** Another motive for mergers turns upon the TBTF doctrine (recall Chapter 11). Since a merger results in a bigger bank, the less likely

22. See Srinivasan (1992). This study examined the before and after performance of all bank mergers completed between 1982 and 1986. The average acquirer had $5,895 billion in assets and the average target had $988.21 million in assets.

23. See Berger and Humphrey (1992).

24. See Srinivasan (1992).

25. See Ersel (2006).

26. See Hunter and Timme (1989).

27. This observation is due to Gorton and Rosen (1992).

28. See Bliss and Rosen (2001).

29. Rose (1987) focused on the performance of national bank mergers during 1976–80. Acquiring banks were found to have lower operating efficiency and productivity than nonmerging banks and their profitability did not increase following the mergers.

30. See Ramakrishnan and Thakor (1991). For empirical evidence that diversification helps, see Eisenbeis, Harris, and Lakonishok (1984).

31. See Berger and Humphrey (1992).

it is that the merged entity will be allowed to fail by regulators. During the rash of bank failures at the beginning of the decade, the U.S. Treasury released a list of large banks that would not be permitted to fail, but the umbrella was expanded. For example, the Bank of New England was protected even though it did not make the list. Such forbearance enhances the appeal of riskier investment strategies. In a sense, TBTF increases the value of the deposit insurance put option (Chapter 10). Hence, a merger of two large banks may increase the value of the governmental safety net to those banks. However, as pointed out in Chapter 12, FDICIA has curtailed the application of TBTF, so that this should be a less important merger motive now.

- **Opportunity to Enter New Markets, Cross-Sell Products and Diversify:** As indicated in previous chapters, banks have been steadily losing market share to nonbank competitors, including the capital market. Money-center banks as well as super-regionals have lost many of their large corporate clients who prefer to borrow funds directly from the capital market by issuing commercial paper and long-term debt. To replace these lost customers, banks have turned to middle-market borrowers. The money centers are thus forced to compete directly with the regional banks, which have long-standing relationships with middle-market firms. Consequently, banks are always looking for new opportunities to grow and these often come via acquisitions. For example, Nations Bank acquired Boatmen's Bankshares in St. Louis to gain access to a key part of the consumer banking market in the Midwestern U.S., and later merged with Bank of America. Citicorp merged with an insurance company, Travelers, in order to be able to cross-sell insurance and banking products. What was interesting about this merger is that it occurred *before* the Glass-Steagall Act was repealed, and hence was technically not legal. Regulators permitted it with the stipulation that the merger would have to be undone if the law was not changed within 5 years!

  Entering new markets also provides diversification opportunities and geographical diversification. Even though the bank's shareholders can diversify on their own, the bank may benefit because the ability to operate in new markets may improve profits in existing markets. Most of these benefits can be expected to accrue to banks of modest size. However, larger banks also may benefit if they have been geographically constrained. Geographical constraints are less binding for asset diversification than for diversifying retail funding sources, but they are present nonetheless. Thus, a merger may be a useful tool to diversify. Evidence indicates that banks bid more for merger partners that offer greater risk-reduction opportunities.[32]

- **Managerial Incentives to Merge for Private Beneitfits:** Even if a merger does not produce benefits for the banks' shareholders, it does result in a bigger bank than either of the merging banks. Since the CEO of the merged bank is typically one of the CEOs of the merging banks, he/she is now at the helm of a bigger institution. This can have many private benefits for the CEO, such as enhanced social prestige, higher perquisites consumption and higher compensation.[33]

32. See Benston, Hunter, and Wall (1992).

33. Bliss and Rosen (2001) show that higher compensation is an important motivation in bank mergers. The seminal paper on private control benefits is Aghion and Bolton (1992).

Hostile Takeovers in Banking

In this section, we discuss evidence of agency problems in banking as well as capital market reactions to takeover attempts.

Agency Problems in Banking

Despite agency problems owing to separation of ownership and control, takeovers in banking were infrequent until recently. This was due to regulatory restrictions and attitudes. The consequent persistence of agency problems and managerial inefficiency led to chronic excess capacity, excessive costs, and poor risk management. This problem was exacerbated by the fact that the banking industry has been buffeted by many changes like deregulation, global competition, technological change, and securitization. These changes required adaptive strategies that many banks' managements could not implement. Unfortunately, many banks are still being managed in an environment of conflicts among owners, managers, and public regulators, and on principles that inhibit adapting to environmental changes. This creates incentives for the more efficiently managed to purchase the less efficiently managed banks.

A lack of managerial adaptation to changing environmental conditions can show up in various ways, most notably in excessive conservatism with respect to risk-taking. An empirical study validated this hypothesis.[34] Comparing *managerially controlled* banks (defined as those whose managers held a small proportion of the bank's stock and were, therefore, more likely to act in a manner not consonant with the shareholders' interests) with *stockholder-controlled* banks (defined as those whose managers held a large portion of the bank's stock and were more likely to act in the stockholders' best interest), the study found that stockholder-controlled banks exhibited significantly higher risk-taking behavior than managerially controlled banks. This provides one indication of how separation of ownership and control can distort bank behavior.

Are Takeovers in Banking a Good Idea?

Despite formidable evidence that takeovers can increase shareholder value, these acquisitions are still criticized, most notably by bank managers. Three main arguments are offered.

(i) Takeovers Bias Managerial Decisions in Favor of Shareholders: Takeovers force management to reflect the interests of their shareholders at the expense of other stakeholders. Consequently, employees, depositors, borrowers, communities, and deposit insurers (taxpayers) suffer. This argument has three major flaws. First, it fails to consider the relative priority of stakeholder rights. Shareholders provide the capital that is most at risk and are, therefore, given certain control rights. These rights empower them to appoint and replace managers. Alternative stakeholders in banks

34. See Saunders, Strock, and Travlos (1990).

have agents other than managers to represent their interests. For instance, depositors are represented by the FDIC, and low-income borrowers by community groups as well as the government. It is the responsibility of management to maximize shareholder wealth, subject to the constraint that the bank's contractual obligations to other groups are honored. In particular, this constraint requires abiding by the wishes of regulators who have "cease and desist" powers over banks. Second, the impact of takeovers on alternative stakeholders is not necessarily negative. Increased efficiency may come at the expense of employee layoffs, but the layoffs would probably have occurred anyway. A bank bloated with excess capacity will fail eventually under the pressure of competition, in which case employees will be laid off anyway.

(ii) Takeovers Induce Myopic Decision Making: Fear of takeovers allegedly predisposes management toward the short run. This focus on quarterly earnings and day-to-day stock price movements may work to the detriment of owners' long-term interests.[35]

While this argument has merit, it is not clear that it is particularly relevant for the debate about whether takeovers are good or bad for *banks*. That is, if the threat of takeovers causes managers to adopt myopic investment policies,[36] then the design of managerial compensation packages should be adapted accordingly. Giving managers "golden parachutes," for example, would lessen their concern with takeovers. This is a double-edged sword, however, since agency problems may worsen if the manager becomes indifferent to takeover threats. Thus, a balance must be struck in designing the golden parachute. The point is that there are ways in which the disciplinary and management efficiency implications of takeovers can be preserved without compromising the efficiency of the bank's long-run investment policies.

(iii) Takeovers Harm Banking Relationships: Finally, some critics argue that takeovers are bad for banks because the adverse publicity harms banking relationships with customers. In particular, the uncertainties raised by a hostile offer may cause some customers to flee to other banks.

This is a reasonable argument. However, it is not in the best interests of the acquiring bank to jeopardize existing relationships at the target bank. Indeed, the acquirer's takeover attempt is likely to be motivated by the prospect of access to the target bank's customer base. In the final analysis, if the surviving bank is more efficiently managed, its long-term survival prospects will be enhanced, and the likelihood of relationships being maintained will be improved. A caveat to this is that a takeover struggle may dissipate proprietary information and human capital, in which case there is a real cost associated with takeovers.

35. Stein (1988) has provided a model in which takeover threats induce myopic investment behavior by managers. The basic idea is that there is asymmetric information between the manager and the capital market. The manager knows the firm's true value, whereas the market attempts to infer this value from current earnings. By diverting cash flows from long-term investments, the manager can inflate current earnings to convey "good news." Even though the market correctly anticipates such behavior, there is a Nash equilibrium in which takeover threats lead to managerial myopia.

36. Thakor (1990, 1993) shows that investment myopia can occur even without takeover threats, given that the manager's objective is to maximize the wealth of current shareholders.

Methods by Which Managers in Target Banks Resist Takeovers

In summary, then, takeovers may exacerbate investment myopia that may be difficult to correct with managerial compensation contracts and also may threaten proprietary information and human capital. While the empirical significance of these costs in banking is unknown, banks have begun to adopt a wide variety of takeover defenses. The five major defensive strategies include the following:[37]

(i) Charter and By-Law Provisions: This strategy impedes, but does not preclude, takeovers. These provisions usually take the form of "Fair Price Provisions" and "Staggered Board Provisions." Fair price provisions mandate that shareholders receive equivalent consideration at both ends of a two-tier bid.[38] Staggered board provisions ensure continuity by having multiyear terms for directors expire at staggered dates. An acquirer thus must wait a year or more to gain a majority of board seats.

(ii) Defensive By-Law Provisions: These consist of three types of provisions: "notice of business and nominations," "action by written consent," and "protection against conflicts of interest in proxy fights." Notice of business and nominations provision generally sets a date well in advance of a board meeting by which a shareholder advises the bank of his/her intent to seek action at a meeting and stipulates the action desired.

The second type of by-law, action by written consent, requires the following:

- Prompt notice of any action taken without written consent of all the stockholders must be given to nonconsenting stockholders.
- Upon receipt of the proper notice from a shareholder who desires shareholder action by written consent, the board must set a record date 10 days after notice is received.
- The record date must be announced promptly and publicly.
- Shareholders must be given at least 20 days after the record date during which they have the option to revoke their consent.
- Consents are valid for a maximum of 60 days after the record date.

These provisions impede takeovers by precluding spontaneous actions, giving dissenters time to marshal defensive efforts.

The third type of defensive by-law provision is designed to ensure that only disinterested directors—those with no conflicts of interest—vote on the takeover, so that shareholders are protected against self-serving behavior by the directors.

(iii) Rights Plans: These are often called "poison pills" or "shark repellant clauses," and remain among the most effective takeover deterrents. Although poison pills come in a variety of forms, the most common is a *share repurchase rights plan*. In a popular version of this plan, holders of the rights,

37. See Herlihy and Shrock (1989) for more details.

38. Two-tier bids and front-end loaded tender offers involve the would-be acquirer offering one price for about 50 percent of the target's outstanding common stock and then using the majority or control thus obtained to "squeeze out" the remaining shareholders in a lower-priced "back-end" merger.

other than the acquirer, are entitled to "flip in" and purchase additional shares of stock in the target bank at a steep discount from market value. The entitlement is triggered by the acquiring crossing a percentage threshold of ownership in the target.[39] Until someone acquires or makes an offer for a specified percentage of the target bank's common stock, the rights are not exercisable.

(iv) "White Squire" and "White Knight" Arrangements: Sometimes a target bank's management seeks protection by turning to another potential acquirer called a "white squire" or a "white knight" in the jargon of Wall Street. Typically, the white knight is first given the opportunity to purchase securities in the target so as to impede the hostile acquirer's ability to acquire the securities needed to gain control. Then the white knight and the target enter into a "stand-still" agreement delineating various restrictions on the white knight's ability to vote its securities and acquire additional voting securities. White knights are typically given a "consolation prize" in case their bid for control fails. This consolation prize deters the hostile acquirer.

(v) Restructuring Defenses: For nonfinancial corporations, corporate restructurings have proved to be an effective defense against takeover attempts, especially those that are intended to sell off pieces of the target after the acquisition. The basic idea is for the target to anticipate which parts of its business an acquirer is likely to divest and to take these actions at its own initiative, sometimes even before a hostile bid has been launched. Often the divested parts are the ones that are losing money. Other restructurings alter the capital structure of the target by substituting debt and/or preferred stock for common stock or through common stock repurchases. The idea is to increase financial leverage before the takeover, so that there is less value to be captured by an acquirer seeking to increase the target's leverage after the takeover for the purpose of enjoying a higher debt tax shield.

Some of the more effective restructuring defenses are unavailable to BHCs owing to government regulations. For example, capital adequacy guidelines limit the ability of a BHC to substitute debt for equity. Of course, most banks are more leveraged than nonfinancial companies, and typically push financial leverage close to the limits of governmental standards. Restructurings that are within reach for most banks typically include elimination of excess capacity and improvements in operating efficiency.

Strategies of Hostile Bidders

To counter the defensive initiatives of potential targets, bidders have an arsenal of weapons at their disposal. These do not necessarily deter defensive measures, but they often make it profitable to attempt takeovers despite the defensive measures of targets. We discuss some of these strategies here.[40]

39. Since the rights are "out of money," they are not included in outstanding shares for purposes calculating earnings and book value per share.

40. See Wolff III (1989).

- The bidder can affect a "creeping tender offer" by taking over a target through discrete open-market purchases of stock.
- The bidder can take a "free ride" by secretly acquiring a 5 percent position and then putting the target in play by tendering for the rest of the target. The eventual success of this strategy depends on obtaining the necessary Federal Reserve approval for the acquisition, but as long as this probability is sufficiently high, the acquirer can profit *ex ante* from trading (on the initial 5 percent ownership) even though takeover defenses by the target may result in the takeover price being so high that the acquirer makes no profit on the shares acquired beyond the initial 5 percent. This encourages takeover bids even in the face of defenses.
- The bidder can initiate a tender offer that induces speculators to purchase large blocks of shares and can then gain control by purchasing shares from the speculators.
- The bidder can seek greenmail—offer to sell back his/her shares to the firm at a negotiated price in excess of market. Management may comply in order to terminate the takeover threat.
- The bidder may initiate a proxy fight.
- The bidder may attempt a "bear hug" of the target. This tactic, which is becoming increasingly popular in banking, involves the bidder sending an unsolicited offer letter to the target's board of directors. Unlike a formal offer, which requires disclosure through a filing with the SEC, the letter usually contains only a vague outline of terms. The idea is for the bidder to elicit a response from a board that may be less than completely satisfied with management. There are three types of bear hugs.[41] The "teddy bear hug" is the most common and the least threatening to management. It presents the general terms of a prospective deal, along with a pledge not to pursue the transaction without the board's approval or go public with an offer. The "brown bear hug" is more aggressive and contains no promises to respect the board's wishes, and veiled threats that the acquirer will enter the target's market anyway if its initial overture if rebuffed. And finally there is the "grizzly bear hug," which has a hostile tone and full public disclosure that an offer has been made. While there have been numerous bear hugs in banking, apparently the only successful grizzly bear hug was Bank of New York's takeover of Irving Bank Corporation in October 1988. Others have tried and failed, including National City Corp.'s 1991 hostile bid for another Cleveland bank—Ameritrust Corp.[42]

Apart from differences in the regulatory environments of banks and nonfinancial corporations, there is not that much difference between banks and other corporations insofar as corporate control contests are concerned. As in the case of nonfinancial corporations, the focus of future corporate control contests in banking is likely to be an attempt to make banks more focused and competitive in the global marketplace.

41. In 2005, New York–based E-Trade Financial Corp. made a grizzly bear hug offer for a rival online brokerage firm, Ameritrade Corp. based in Omaha, Nebraska, prompting Ameritrade to acquire the U.S. online brokerage operations of Toronto-based TD Bank Financial Group.

42. See Engen (2005).

Price Reactions to Mergers and Takeovers: The Empirical Evidence

Considerable scientific evidence has been gathered on how the stock market reacts to corporate control contests in banking. The studies that focus on short-term stock performance find that acquirers on average experience negative announcement returns while targets gain. Moreover, these studies fail to uncover total gains from consolidation. Few studies examine the long-run stock return performance of acquiring banks. Those that do find that acquirers significantly underperform their peers in the long run (2 to 3 years after the acquisition).[43] Studies of mergers and acquisitions in Europe encounter similar results.

Conclusion

Consolidations have been changing the map of American and European banking for over two decades. New alliances are being forged and new opportunities exploited. The merger of Nations Bank with Bank of America created perhaps the only truly national retail banking franchise in the U.S. And the merger of Citicorp and Travelers created Citigroup, which had $1.5 trillion in assets in 2005 and was the largest bank in the U.S. at the time. Of course, banking is not unique in this respect. Many other industries, such as automobiles, airlines and telecommunications, are all in the same state of structural flux.

This wave of mergers has coincided with the highest increase in productivity in the financial sector in recent years. It is hard to tell what portion of these productivity gains were attributable to the consolidations since there were numerous other factors at work. For instance, competitive pressures have eliminated waste, shedding of redundant capacity, repricing of services, and better management. Moreover, regulators have expanded bank powers, and the low and positively sloped yield curve enabled banks to profit from duration transformation. Nonetheless, mergers and acquisitions are expected to be a dominant part of the global banking landscape for many years.

Review Questions

1. Would you expect to see more or less takeovers in banking than in other industries? Explain your answer.
2. What are the principal motivations for banking mergers? Discuss the empirical evidence in connection with the *reasons* why banks merge. Does the evidence indicate significant synergies that would warrant mergers of the magnitudes we have witnessed?
3. How is corporate governance in banking affected by takeovers?
4. Discuss the theory and empirical evidence related to the effects of mergers and acquisitions on the wealths of target and acquiring banks' shareholders.
5. What is the role of bank regulation in the mergers and acquisitions process?

43. See Piloff and Santomero (1998) and Rhoades (1994, 2000) for overviews of these findings.

6. Discuss takeover defenses and whether these benefit the shareholders of the target bank.
7. Why would we expect the stock price reaction to a takeover attempt to exceed that to a proposed merger? Be sure to discuss both the theoretical and the empirical underpinnings of differential announcement effects.
8. What is "limit pricing," and what is its relevance in bank mergers?
9. If you were a regulator, what would your attitude be toward mergers in banking? Explain your view.

References

Aghion, Phillipe, and Patrick Bolton, "An Incomplete Contracts' Approach to Financial Contracting," *Review of Economic Studies* 59, 1992, 473–494.

Benston, George J., William C. Hunter, and Larry D. Wall, "Motivations for Bank Mergers and Acquisitions: Enhancing the Deposit Insurance Put Option Versus Increasing Operating Net Cash Flow," Federal Reserve Bank of Atlanta Working Paper 92–4, April 1992.

Berger, Allen N., and David B. Humphrey, "Megamergers in Banking and the Use of Cost Efficiency as an Antitrust Defense," *The Antitrust Bulletin* 33, Fall 1992.

Bliss, Richard, and Richard Rosen, "CEO Compensation and Bank Mergers," *Journal of Financial Economics* 61–1, 2001, 107–138.

Boot, Arnoud W.A., "Why Hang on to Losers: Divestitures and Takeovers," *The Journal of Finance* 47–2, September 1992, 1401–1424.

Boot, Arnoud, Todd Milbourn and Anjan V. Thakor, "Megamergers and Expanded Scope: Theories of Bank Size and Activity Diversity," *Journal of Banking and Finance* 23, 1999, 195–214.

Cotter, James F., and Marc Zenner, "How Managerial Wealth Affects the Tender Offer Process," *Journal of Financial Economics*, 35, 1994, 63–97.

Cover Story, "The Age of Consolidation," *Business Week*, October 14, 1991.

Crockett, Barton, "Tech Bosses' Everest: Making Mergers Work," *American Banker*, June 17, 1992.

Desai, Anand, and Roger Stover, "Bank Holding Company Acquisitions: Stockholder Returns and Regulatory Uncertainty," *Journal of Financial Research* 8, 1985, 145–156.

Eisenbeis, Robert A., Robert S. Harris, and Josef Lakonishok, "Benefits of Bank Diversification: The Evidence from Shareholder Returns," *The Journal of Finance* 39–3, July 1984, 881–894.

Eisenbeis, Robert, "Bank Holding Companies and Public Policy," in *Financial Services: The Changing Institutions and Government Policy* (George Benston, ed.), Prentice-Hall, Inc., Englewood Cliffs, N.J., 1983, 140–143.

Engen, John R., "Many Happy Returns," *Bank Director*, 4th Quarter 2005, 26–28.

Engen, John, "The Bear Country Survival Guide," *Bank Director*, 3rd Quarter 2005, 48–55.

Ersel, Isil, "The Effect of Bank Mergers on Loan Prices: Evidence from the U.S.," Working Paper, The Ohio State University, September 2006.

Gorton, Gary, and Richard Rosen, "Overcapacity and Exit From Banking," Working Paper, Wharton School, University of Pennsylvania, June 1992.

Grossman, Sanford J., and Oliver D. Hart, "Takeover Bids, the Free-Rider Problem and the Theory of the Corporation," *Bell Journal of Economics* 11, 1980, 42–64.

Hawke, John D., Jr., "The Changing Environment for Takeovers," *Issues in Bank Regulation*, Winter 1989, 5–8.

Herlihy, Edward, D., and Gregory J. Shrock, "Defending Against Hostile Takeovers," *Issues in Bank Regulation*, Winter 1989, 9–14.

Hirshleifer, David, and Anjan V. Thakor, "Corporate Control Through Board Dismissals and Takeovers," Working Paper, UCLA and Indiana University, 1993.

——, "Managerial Conservatism, Project Choice and Debt," *The Review of Financial Studies* 5–3, 1992, 437–470.

——, "Managerial Performance, Boards of Directors and Takeover Bidding," Working Paper, UCLA and Indiana University, 1994, forthcoming, *Journal of Corporate Finance*.

Huber, Stephen K., *Bank Officer's Handbook of Government Regulation*, Second Edition: 1993 Cumulative Supplement, Warren Gorham Lamont, Boston, 1993.

Humphrey, David B., "Why Do Estimates of Bank Scale Economies Differ?" *Federal Reserve Bank of Richmond Economic Review*, September/October 1990, 38–50.

Hunter, William C., and Stephen G. Timme, "Does Multiproduct Production in Large Banks Reduce Costs?" *Federal Reserve Bank of Atlanta Business Review*, May/June 1989, 2–11.

James, Christopher, and Peggy Wier, "An Analysis of FDIC Failed Bank Auctions," *Journal of Monetary Economics* 20, 1987a, 141–153.

——, "Returns to Acquirers and Competition in the Acquisition Market: The Case of Banking," *Journal of Political Economy* 95, 1987b, 355–370.

"Japanese Corporate Governance: Dawning," *The Economist*, May 22, 1993.

Jensen, Michael C., "Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers," *American Economic Review* 76, 1986, 323–329.

Jensen, Michael C., and Jerold B. Warner, "The Distribution of Power Among Corporate Managers, Shareholders, and Directors," *Journal of Financial Economics* 20, 1988, 3–24.

Jensen, Michael C., and William H. Meckling, "Theory of the Firm: Managerial Behavior, Agency Cost and Ownership Structure," *Journal of Financial Economics* 3, 1976, 305–360.

Kantrow, Yvette D., "Why Chemical's Name Will Be the Survivor," *American Banker*, July 17, 1991.

Lash, Nicholas A., *Banking Laws and Regulations: An Economic Perspective*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1987.

Layne, Richard, "'Bear Hugs' On the Rise as Takeover Tactic," *American Banker*, July 30, 1993.

Martin, Kenneth J., and John J. McConnell, "Corporate Performance, Corporate Takeovers, and Management Turnover," Working Paper, University of Iowa, Iowa City, 1989.

Matthews, Gordon, "Mergers and Acquisitions Hit a Record $22.5 Billion in '93," *American Banker*, January 11, 1994.

Mester, Loretta J., "Owners Versus Managers: Who Controls the Bank?" *Federal Reserve Bank of Philadelphia Business Review*, May/June 1989, 13–23.

Milbourn, Todd, Richard Shockley, and Anjan V. Thakor, "Managerial Career Concerns and Investments in Information," *RAND Journal of Economics* 32–2, 2001, 82–99.

Milligan, Jack, "Does M&A Still Matter?" *Bank Director*, 3rd Quarter, 2005, 28–34.

Pilloff, Steven J., and Anthony M. Santomero, "The Value Effects of Bank Mergers and Acquisitions, in: Bank Mergers and Acquisitions," Yakov Amihud and Geoffrey Miller (eds.), Boston: Kluwer Academic Publishers, 1998.

Ramakrishnan, Ram T.S., and Anjan V. Thakor, "Cooperation Versus Competition in Agency," *Journal of Law, Economics and Organization* 7–2, 1991, 248–283.

Rhoades, Stephen A., "A Summary of Merger Performance Studies in Banking, 1980–1993," and an assessment of the operating performance and event study methodologies, Staff Study 167 – Board of Governors of the Federal Reserve System, 1994.

——, "Bank Mergers and Banking Structure in the United States," 1980–98, Board of Governors of the Federal Reserve System, Staff Study, 174, 1–32, 2000.

Rose, Peter S., "The Impact of Mergers in Banking: Evidence From a Nationwide Sample of Federally Chartered Banks," *Journal of Economics and Business* 39, 1987, 289–312.

Saunders, Anthony, Elizabeth Strock, and Nickolaos G. Travlos, "Ownership Structure, Deregulation, and Bank Risk Taking," *The Journal of Finance* 45–2, June 1990, 643–654.

Shleifer, Andrei, and Robert Vishny, "Large Shareholders and Corporate Control," *Journal of Political Economy* 94, 1986, 461–488.

Song, Fenghua, and Anjan, V. Thakor, "Information Control, Career Concerns and Corporate Governance," *Journal of Finance*, August 2006, 1845–1896.

Srinivasan, Aruna, "Are There Cost Savings From Bank Mergers?" *Federal Reserve Bank of Atlanta Economic Review*, March/April 1992, 17–28.

Stein, Jeremy, "Takeover Threats and Managerial Myopia," *Journal of Political Economy* 96–1, February 1988.

Stulz, René, "Managerial Control of Voting Rights, Financing Policies and the Market for Corporate Control," *Journal of Financial Economics* 20, 1988, 25–54.

Thakor, Anjan V., "Information, Investment Horizon and Price Reactions," *Journal of Financial and Quantitative Analysis*, December 1993.

——, "Investment Myopia and the Internal Organization of the Capital Investment Decision," *Journal of Law, Economics and Organization* 6, Spring 1990, 129–154.

Wolff, William F., III, "An Investment Banking View of Hostile Takeovers," *Issues in Bank Regulation*, Winter 1989, 23–25.

This page intentionally left blank

CHAPTER ♦ 15

Investment Banking

"I maintain that the Money Market is as concrete and real as anything else; that it can be described in plain words; that it is the writer's fault if what he says is not clear."

Walter Bagehot (1873)

Glossary of Terms

Gramm-Leach-Bliley Act: The 1999 act that dismantled the Glass-Steagall act restrictions separating commercial and investment banking (see Chapter 12).

QAT: Qualitative Asset Transformation (see Chapters 2 and 3).

SPE or SPV: Special Purpose Entity or Special Purpose Vehicle, a structure used for off-balance sheet transactions.

Introduction

With the passage of the Gramm-Leach-Bliley legislation of 1999 and the accompanying dismantling of the Glass-Steagall Act, spatial and functional integration of financial services was encouraged for U.S. financial intermediaries, making the U.S. banking system more like its counterparts in Latin America and Europe. In this chapter, we discuss investment banks. Our discussion will be brief since many of the issues that concern investment banks—derivatives, securitization etc.—have been covered in previous chapters.

We will describe the role of investment banks and the economic services that they provide. This will provide a context for discussing the characteristic contracts investment banks offer their clients.

TABLE 15.1 Bloomberg 20 Top Investment Banks

Company	2004 Total Fees ($ Millions)	Publicly Traded
Citigroup	3,656	Yes
Goldman Sachs	3,605	Yes
Morgan Stanley	3,306	Yes
JPMorgan Chase	2,977	Yes
Merrill Lynch	2,706	Yes
UBS	2,352	Yes
Credit Suisse	2,113	Yes
Deutsche Bank	1,843	Yes
Lehman Brothers	1,607	Yes
Bank of America	981	Yes
ABN Amro Bank	914	Yes
Nomura Securities	744	Yes
RBC Capital Markets	700	No
HSBC	673	Yes
Rothschild	618	No
Daiwa Securities	602	Yes
Lazard	535	No
Wachovia	473	Yes
Bear Stearns	438	Yes
BNP Paribas	424	Yes

Source: Bloomberg Markets; Company Websites.
Note: The Bloomberg 20 ranks investment banks by total fees collected in 2004 for underwriting securities and advising on mergers and acquisitions worldwide.

Investment banking is big business. Revenues surpassed $200 billion in 2000, accounting for 0.6 percent of world GDP,[1] and in the U.S. the securities industry employs 780,000 people or 0.6 percent of total U.S. employment. The top 20 investment banks in the U.S., ranked according to their underwriting and M&A advisory fees, are listed in Table 15.1.

The economic value of the many brokerage and qualitative asset transformation services provided by investment banks is reflected in their revenues. Table 15.2 shows the disclosed fees earned by investment banks globally from 1995–2004 for their securities underwriting activities. As the table shows, the largest firms' revenues grew from $8.5 billion to slightly over $10 billion, while the industry as a whole grew from slightly under $14 billion to over $15 billion in underwriting revenues.

What Investment Banks Do

An investment bank is an FI that specializes in: (i) raising financial (debt and equity) capital; (ii) advising on corporate mergers and acquisitions; and related transactions; (iii) wealth management; (iv) financial and economic research; (v) general financial advisory services; (vi) sales and trading of securities, commodities and currencies; and

1. See Special Feature, *Business World* (Philippines), November 16, 2005.

TABLE 15.2 Global Disclosed Investment Bank Fees From Underwriting

Year	Top 10 Total ($ Millions)	Number of Issues	Industry Total ($ Millions)	Number of Issues
1995	8,503	3,759	13,988	10,196
1996	12,186	4,748	20,327	12,096
1997	13,095	4,758	20,452	10,248
1998	13,829	5,815	20,539	9,810
1999	15,619	6,029	21,855	8,997
2000	16,073	5,489	21,351	7,717
2001	14,217	6,036	18,652	8,526
2002	10,290	5,112	14,628	6,722
2003	9,577	5,855	14,424	8,106
2004	10,088	4,946	15,216	7,049

Source: Thomson SDC Platinum.
Note: Disclosed fees from global debt equity, & equity-related offerings.

(vii) other ancillary activities such as custodial services. Most of the major Wall Street investment banks are active in a wide variety of these activities. Smaller investment banks tend to specialize more narrowly.

We will now describe the services investment banks provide in each of these categories. See Figure 15.1.

(i) Raising Financial Capital: An investment bank can help a firm to raise funds to finance a major investment project, acquire another company, restructure its balance sheet, expand operations, or other business purposes. Capital can include common equity, preferred equity, debt, as well as "hybrid" securities like convertible debt or

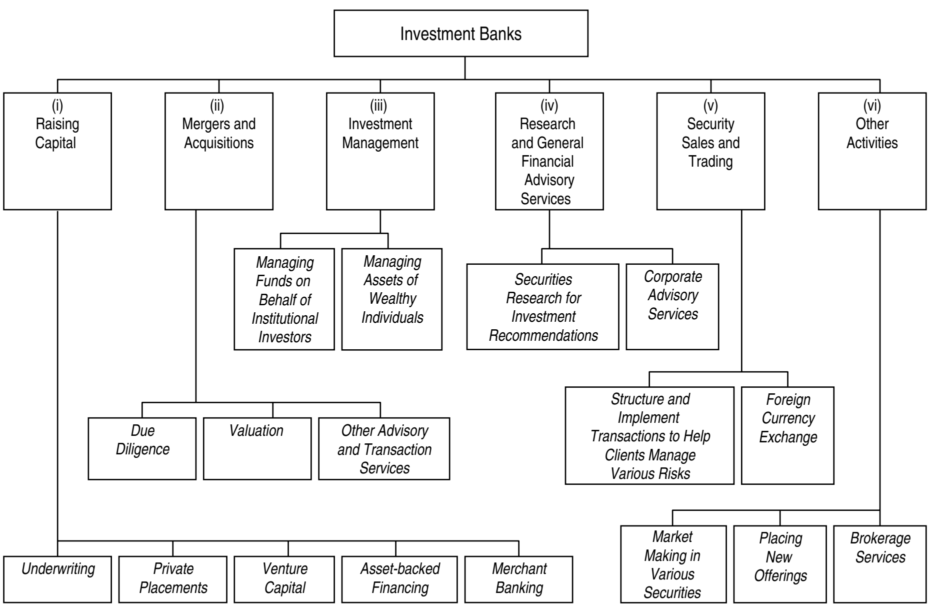


FIGURE 15.1 Services Provided by Investment Banks

debt with warrants attached. Investment banks possess specialized knowledge as to how to structure transactions to meet their clients' specific objectives. Investment banks differ from commercial banks in that investment banks have no access to governmentally insured deposits or to LLR facilities of the Federal Reserve. They are, therefore, not subject to the same regulations as commercial banks. However, the Securities and Exchange Commission monitors and regulates the activities of investment banks in the U.S.

Investment banks raise capital through underwriting, private placements, venture capital, asset-based financing, and merchant banking (see Figure 15.1). We briefly describe each below.

Underwriting: Investment banks verify financial data and business claims, facilitate pricing of claims, and perform due diligence. Most offerings are "firm commitment" underwritings in which investment banks effectively purchase securities from the issuer for resale to the public. In the case of equities, they do this through Initial Public Offerings (IPOs) as well as secondary offerings. They also advise on debt issues to the public markets. Investment banks arranged over half of the total financing provided to U.S. nonfinancial businesses in 2001.[2] A breakdown of total underwriting dollar volumes for U.S. companies by securities is provided in Figure 15.2. Investment banks helped underwrite $4.1 trillion of debt and equity securities worldwide in 2004. Global equity and debt underwriting data are provided in Table 15.3.

IPO underwriting has been a lucrative business for investment banks; but their profits are being challenged. The underpricing of IPOs—the price at which the average IPO is sold is typically lower than the price at the end of the first day of

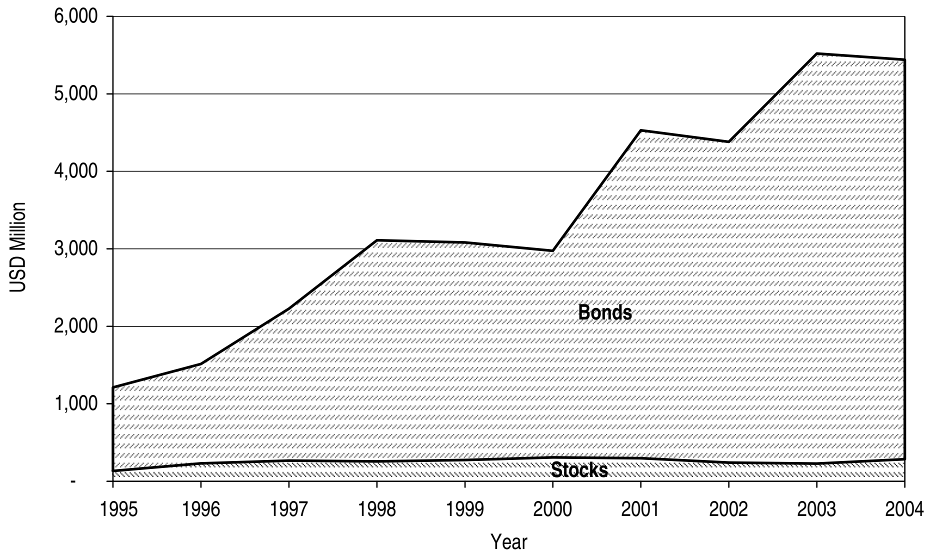


FIGURE 15.2 Breakdown of Total Underwritings for U.S. Companies
Source: Thomson Financial.

2. See United States General Accounting Office (2003).

TABLE 15.3 Total Global Equity and Debt Underwritings

	All Securities		Equity		Debt	
Year	Dollar Amount ($ Millions)	Number	Dollar Amount ($ Millions)	Number	Dollar Amount ($ Millions)	Number
1995	1,663,388	16,204	219,063	4,451	1,444,325	11,754
1996	2,230,329	20,390	322,130	5,533	1,908,199	14,858
1997	2,510,220	19,295	395,087	4,165	2,115,133	15,133
1998	3,046,717	16,535	372,847	3,044	2,673,870	13,501
1999	3,222,521	15,128	492,117	3,415	2,730,404	11,716
2000	3,144,795	16,191	606,667	4,424	2,538,128	11,774
2001	3,563,714	16,941	344,996	3,126	3,218,718	13,823
2002	2,861,217	14,745	280,974	2,785	2,580,242	11,964
2003	3,684,125	16,084	297,156	3,452	3,386,969	12,640
2004	4,105,791	17,209	477,169	4,599	3,628,622	12,621

Source: Thomson SDC Platinum.
Note: Equity includes common stock, convertible preferred stock, and nonconvertible preferred stock. Debt includes convertible debt and nonconvertible debt.

trading—has been much publicized,[3] and this has led to circumventing initiatives like Dutch auctions (see Appendix 15.1 for a discussion of various theories of IPO underpricing and Dutch auctions). In fact, Google's 2004 decision to completely bypass investment bankers in favor of an Internet-based Dutch auction-IPO was seen by many as the possible death knell for IPO underwriting profits for investment banks. This has not yet happened, but the threat of it happening in the future cannot be dismissed. Another development that has negatively affected investment bank profitability in this business is the exposure that there were allegations of conflicts of interest between the underwriting and security research arms of investment banks. That is, the security analysts' opinions of the values of securities were allegedly influenced by underwriting relationships that their employers had with the firms that had issued these securities. New York State Attorney General Eliot Spitzer's investigation, culminating in 2002, accused investment banks of widespread conflicts of interest and pervasive unfair trading. Spitzer's investigation claimed to show that equity analysts' compensation was tied to their ability to attract big IPO clients, which provided incentives for analysts to inflate valuations. It also claimed that investment banks were attracting clients by promising CEOs preferential allocations of shares in underpriced IPOs.[4]

These allegations and the accompanying threat of enforcement actions persuaded ten major investment banks to agree to the so-called global settlement in 2003, whereby they paid sizeable fines. New rules were formulated that prohibited investment bankers and analysts from collaborating and also altered the terms of IPO allocations. These developments put pressure on the profit margins of investment banks.

3. See, e.g., Jay Ritter and Ivo Welch, "A Review of IPO Activity, Pricing, and Allocations," *Journal of Finance* 57–4, 2002, pp. 1795–1828.
4. See Garver (2005) for a discussion of this issue.

Private Placements: As an alternative to a public offering, the banker may distribute newly issued or newly available debt or equity claims to a small number of larger, typically institutional buyers. These private placements are typically less costly to distribute because SEC registration requirements are less stringent or the supposition that the buyers are more sophisticated. In addition, the securities may not be traceable. These transactions are more like brokerage undertakings rather than QAT in that the purchasers are buying directly from the issuers.[5] The banker may never own the securities in question and the banker takes a fee, but no spread. Private placements are more common with debt than equity in that companies typically prefer a wide distribution of equity. Wide distribution of debt is likewise preferred, but is rarely feasible.

Venture Capital: Investment banks also provide capital and strategic guidance to younger and smaller companies and may manage venture capital pools or even invest their own capital. Venture capital has grown significantly in volume in the past 10 years, peaking in 2000. Table 15.4 provides data on U.S. venture capital investments during 1995–2004. Venture capitalists normally accept greater risks when they provide financing because they often fund innovative ventures with prospects based on little or no experience. To be compensated for this risk, they demand high expected returns. This is reflected in Table 15.5, which shows data on U.S. venture capital returns during 1995–2004. The high volatility in returns reflects the high risk.

Securitization and Asset-Based Financing: Investment banks help their clients use their existing assets to obtain additional financing without actually having to sell off these assets. The process by which this is done is called securitization, and the securities that are created in the process are called asset-backed securities. For example, a company might have (uncollected) receivables. It could then issue securities that permit the buyers of these securities to receive cash flows as these receivables are collected. The receivables are segregated from the rest of the company's assets for

TABLE 15.4 U.S. Venture Capital Investments

Year	Dollar Amount ($ Millions)	Number of Deals
1995	7,879	1,773
1996	11,014	2,471
1997	14,612	3,084
1998	20,811	3,553
1999	53,476	5,396
2000	104,701	7,809
2001	40,703	4,456
2002	21,698	3,057
2003	19,585	2,865
2004	21,635	2,966

Source: PricewaterhouseCoopers/Thomson Financial Venture Economics/National Venture Capital Association MoneyTree Survey.

5. This allows them to take advantage of intertemporal and cross-sectional informational reusability, i.e., the ability to use the same information through time and across clients, having invested only once in acquiring this information. See Bhattacharya and Thakor (1993).

TABLE 15.5 U.S. Venture Capital Returns

Year	Annual Return
1995	35.93
1996	52.58
1997	41.04
1998	14.74
1999	133.24
2000	209.41
2001	(46.87)
2002	(30.27)
2003	(11.33)
2004	10.20

Source: Cambridge Associates LLC U.S. Venture Capital Index.
Note: Annual return based on data compiled from 1,060 U.S. venture capital funds, including fully liquidated partnerships. All returns are net of fees, expenses, and carried interest. Yearly returns ended September 30th.

the purpose of delineating them as specific assets to be used to back up investors' claims. Financing is raised by the issuing company due to the money investors pay to purchase the securities issued against the receivables. Recall we discussed securitization in Chapter 9.

Merchant Banking: These activities involve the investment bank committing its own capital to facilitate a variety of client transactions. That is, these are transactions involving qualitative asset transformation. Merchant banking transactions may include loan commitments, syndicated loans, highly leveraged transactions, bridge loans, and so on. A loan commitment is a promise by the bank to make a loan available in the future for a preidentified purpose such as an acquisition or a major project. A syndicated loan is one in which the bank is a member of a group making a loan to a borrower. The syndicated loan market is a very large one. Table 15.6 provides data on U.S. syndicated loans, which were discussed in earlier chapters. A highly leveraged transaction is one in which the bank loan is part of a financing package for an acquisition or some other form of asset investment and the financing package involves a relatively high debt-equity ratio.[6] A bridge loan is a temporary loan that serves as a bridge to more permanent future financings. For example, an investment bank may extend a bridge loan to help finance an acquisition, with the idea being that at some point in the near future the client will issue bonds to repay the bridge loan, and it is not uncommon for the bank that extended the bridge loan to be asked to underwrite the bond issue.

In their role as providers of finance, investment (as well as commercial) banks make extensive use of credit derivatives, both as buyers and sellers. These are designed to protect the bank against default by the borrower since the seller of the

6. A transaction in which credit is extended in connection with leveraged buyouts, mergers and acquisitions, or corporate restructuring, and where the credit results in a total debt/asset ratio exceeding 75 percent is called a "highly leveraged transaction." See Highly Leveraged Transactions (HLTs): *Office of the Director of the Division of Bank Supervision of the Federal Deposit Insurance Corporation*, BL-21-89, May 10, 1989.

TABLE 15.6 U.S. Syndicated Loans

Year	Dollar Amount ($ Millions)
1999	676,851
2000	1,039,738
2001	993,927
2002	880,510
2003	818,055
2004	1,290,841

Source: Bloomberg Custom League Tables.

TABLE 15.7 U.S. Derivatives Contracts Notional Amount Outstanding ($ Billions)

Year	Interest Rate	Foreign Exchange	Other Derivatives	Credit Derivatives	Total
1995	11,095	5,387	378		16,861
1996	13,427	6,241	367		20,035
1997	17,085	7,430	494	55	25,064
1998	24,785	7,386	684	144	32,999
1999	27,772	5,915	843	287	34,817
2000	32,938	6,099	1,080	426	40,543
2001	38,305	5,736	950	395	45,386
2002	48,347	6,076	1,016	635	56,074
2003	61,856	7,182	1,043	1,001	71,082
2004	75,518	8,607	1,409	2,347	87,880

Source: U.S. Office of the Comptroller of the Currency, Bank Derivatives Report, Third Quarter 2005.
Note: Numbers may not add due to rounding.

credit derivative to the bank is obligated to pay the bank if the borrower defaults. Use of credit swaps has become a part of prudent risk management rather than an assessment by the bank that the borrower *will* default. Credit derivatives permit credit risk to be spread across various capital market participants in order to diminish credit-risk concentrations. The global use of credit derivatives has grown significantly in the past 2 decades, as part of the explosive growth in the overall derivatives market. Table 15.7 provides data on U.S. derivatives contracts and Table 15.8 indicates the size of the global swaps and derivatives market.

(ii) Mergers and Acquisitions: Investment banks provide a variety of services to help their clients with mergers and acquisitions: (a) due diligence; (b) valuation; and (c) other advisory and transaction services. We describe each briefly below.

Due Diligence: When a company is considering the acquisition of another company, it needs to examine the target company's market and financial condition to ensure that it does not end up acquiring unforeseen problems or overpaying. The process by which the information relevant to this is collected and analyzed is called due diligence, and investment banks possess expertise in providing this service.

Valuation: Any time a company is considering acquiring another company, it needs to establish the maximum price it is willing to pay. This is a blend of both the science

TABLE 15.8 Global Swaps and Derivatives Notional Amount Outstanding ($ Billions)

Year	Interest-Rate Swaps, Interest-Rate Options, and Currency Swaps	Credit Default Swaps	Equity Derivatives
1995	117,713		
1996	25,453		
1997	29,035		
1998	50,997		
1999	58,265		
2000	63,009		
2001	69,207	919	
2002	101,318	2,192	2,455
2003	142,307	3,779	3,444
2004	183,583	8,422	4,151

Source: International Swaps and Derivatives Association, 2005.
Note: ISDA began surveying for credit default swaps in 2001 and equity derivatives in 2002.

of finance and an art form in evaluating information and making the right assumptions. The task is especially challenging for private companies where no market price is available as a benchmark, but also in the case of public companies since possible synergies due to the merger would not be reflected in the target's preacquisition stock price and would need to be assessed. Investment banks have developed expertise in valuation that they share with their clients.

Other Advisory and Transaction Services: These include advising the client on the best type of transaction, preparing a selling memorandum, participating in negotiations, and assisting the client's board of directors with discharge of its fiduciary duties.

(iii) Investment Management: Investment banks engage in investment management primarily of two types: (a) managing funds on behalf of institutional investors; and (b) managing the assets of wealthy individuals. We describe each briefly below.

Managing Funds on Behalf of Institutional Investors: Investment banks manage mutual funds, hedge funds, unit investment trusts, leveraged buyout firms, and private equity funds. These funds are often managed on behalf of institutional investors like pension funds and life insurance companies in global capital markets.

Managing Assets of Wealthy Individuals: Investment banks also operate private banking accounts in which they provide customized wealth management services to wealthy ("high net worth") individuals. The services include help with investments, retirement planning, insurance and estate planning.

(iv) Research: Investment banks conduct research on companies, financial markets and the economy in order to provide informed, high-quality advice to their clients. This research is sold directly or indirectly as part of a package of services. Banks have a history of providing investment advice to include transactions on which they earn both spreads and fees, and these include both debt and equity securities.

Equity Securities Research for Investment Classifications: Investment banks employ research analysts who are specialists in conducting research on the economic conditions prevailing in the industry and in the overall market, and also in investigating idiosyncratic firm-specific factors so as to be able to value the traded equities of various firms. The research and the resulting valuation then serve as the bases upon which research analysts classify specific securities into buy, sell or hold categories for clients.

Debt Securities Research: Investment banks also employ research analysts to conduct research on the debt instruments of various companies. Analysts involved in this are experts in credit risk assessments.

(v) Corporate Advisory Services: Investment banks provide a host of advisory services to their corporate clients. These advisory services, informed by their research, include help with corporate reorganizations, resulting in recommendations about the sale of specific assets, the issuance of securities, and the possible negotiation of the sale of the entire company. In addition, banks offer advice relating to joint ventures, privatizations, spinoffs, tender and exchange offers, leveraged buyouts, and defense strategies against hostile takeovers.

(vi) Security Sales and Trading: Investment banks are active in the sales and trading of various securities. These include stocks, as well as fixed-income products like government bonds, Eurobonds, money market instruments, swaps, corporate bonds, municipal bonds, asset-backed securities, floating rate notes, mortgage bonds, bond options, and other more exotic derivatives. Investment banks provide sales and trading services in principally three ways: (a) market making; (b) placing new offerings; and (c) brokerage services. We describe each of these briefly below.

Market Making: As a market maker, an investment bank promotes price stability and continuity by holding inventories of the security, with a willingness and ability to step in and redress temporary imbalances in supply and demand. For example, an investment bank may stabilize prices during an IPO and then act as a market maker in the new securities.

Placing New Offerings: Investment banks actively market new securities either as an agent or a principal. This is part of their capital-raising function discussed earlier.

Brokerage Services: Investment banks also engage in sales and trading for institutions as well as individuals. This is an aspect of market making.

(vii) Other Ancillary Activities: This encompasses a variety of other activities of investment banks that do not fall in any of the previous categories. These activities include: (a) the structuring and implementing of transactions to help clients manage various risks, and (b) custodial and corporate trust services. Each is briefly described below.

Structure and Implement Transactions to Help Clients Manage Various Risks: Firms that are clients of investment banks face an assortment of risks in the course of business. Investment banks help their clients manage these risks. Such risk management often involves the use of derivatives and off-balance sheet transactions. An example of how an investment bank might help a client manage risk would be by arranging an interest-rate swap for a firm that has fixed-rate assets and floating-rate liabilities. Such a firm is exposed to the risk that interest rates may unexpectedly rise and push up the cost of refinancing its liabilities above the fixed yield on its assets. An investment bank could help such a firm enter into a swap transaction, which is typically off-balance sheet, whereby this firm could exchange its floating-rate liability payments for the fixed-rate obligation of another firm. The investment bank's roles here are typically to design the swap to meet its client's needs and then to find a counterparty to the swap (i.e. the firm with the fixed-rate obligations that would be interested in entering into a swap with the investment bank's client).

Custodial and Corporate Trust Services: Banks provide custodial and trust services to their corporate clients, as part of the package of investment banking services. These include a variety of fiduciary and agency products to clients, including corporations and government entities. Included in these products are trustee, paying agent and registrar services, successor trustee services, document custodial services, back-up and master servicing, securities administration services, and escrow services.

Separation of Investment Banking Activities

Because investment banks have potential access to significant amounts of proprietary information about their clients, they go to great lengths to specify to their employees rules and procedures to ensure that the bank neither trades on this proprietary information nor does it make buy/sell recommendations based on this information. That is, investment banks erect "Chinese Walls" that separate the banking part of their business from the marketing side of their business.

The Chinese Wall is designed to eliminate the flow of nonpublic information obtained from the investment bank's clients in order to ensure that such information does not advantage another part of the investment bank. Thus, the Research Department is prohibited, in its research reports or buy/sell classifications, from using proprietary (non-public) information that the investment banking part of the bank may have obtained through its investment banking relationships.

Chinese Walls are typically supplemented by other restrictions on employee trading. Investment banks are also typically explicit in specifying in their internal compliance documents the specific circumstances in which employees can scale the Chinese Wall. For example, a research analyst with specialized industry knowledge may be invited to assist with the deal. However, there are strict restrictions on the analyst's ability to use in subsequent research reports or buy/sell/hold recommendations the information acquired in this interaction, as long as it is not public. In fact, the investment bank would typically withhold publication of research on the company during the time that the analyst is involved with the investment banking deal.

There is a good economic reason why investment banks go to such great lengths to make explicit provisions like Chinese Walls and specify related proscriptions in their compliance documents to guide employees about how to avoid conflicts of interest and abuse of privileged information. The reason is reputational capital, which may be even more important than financial capital for investment banks. Consequently, like bond rating agencies and other reputation-based intermediaries, investment banks strive to develop and preserve reputational capital, and reputational risk is one of the risks they attempt to manage.[7]

7. Investment banks typically conduct periodic reputational risk reviews.

Risk Management, Structured Finance, and Investment Banks

As indicated in the previous section, investment banks help their clients manage risk in a variety of ways. One of these ways is to arrange what are called "structured finance" transactions. The term "structured finance" is used to refer to the mix of securities used in structuring "off-balance sheet" transactions, i.e., transactions whose entire value does not show up on the clients' balance sheet (see Chapters 8 and 9). Such structured finance transactions are commonplace in various industries. Structured financing transactions generally isolate the firm's assets and obligations in a "structure" that is apart from the main operations of the sponsor. The structure is typically called a *Special Purpose Entity* (SPE) or *Special Purpose Vehicle* (SPV). Its cost of capital may differ from that of the sponsor's, and the sponsor's control over the structure is generally more limited than in the case of on-balance sheet financings. It is common to contract the management of the SPE or SPV to a trustee who must be independent of the sponsor's. The trustee's discretion, in turn, is limited to certain types of transactions such as mortgage investing, project construction, project leasing, etc. An important reason for setting up a segregated SPE or SPV is to reduce agency or informational costs.[8]

An important economic function of these structured finance transactions is to help firms manage risks and raise project financing with levels of financial leverage that would not be optimal had the financing been structured as routine on-balance sheet debt financing. That is, structured finance is another way for a corporation to raise financing. Just as investment banks help companies raise money for general purposes through equity, preferred stock, commercial paper and a variety of debt securities, these banks also help companies raise money to meet particular financing needs by using structured finance transactions. The structured finance market is quite large and involves the world's major financial firms. Apart from the many examples of structured finance we will discuss later in this report, other examples of structured finance transactions include collateralized debt obligations (such as mortgage-backed securities and credit-card securitization), debt-equity hybrid securities, leases, convertible bonds and convertible preferred stock. There is a great variety of structured finance contracts and the specifics may differ significantly from client to client, but they all are intended to help companies increase liquidity, diversify funding sources and improve risk management.

There is such a large number of off-balance sheet transactions used to manage risks that it would be difficult to deal with them exhaustively. However, conceptually, we can think of these transactions as being designed to help manage the following kinds of risks: price volatility risk of financial investments (such as equity holdings in other companies), commodity price risk, and illiquidity risk associated with specific assets in the firm's portfolio.

These three types of risks are hedged/managed in a variety of ways, and investment banks help their clients come up with effective mechanisms to hedge these risks. For example, the price volatility risk of financial investments can be managed through limited partnerships and trading derivative securities directly in the capital

8. See Shah and Thakor (1987) and Chemmanur and John (1996).

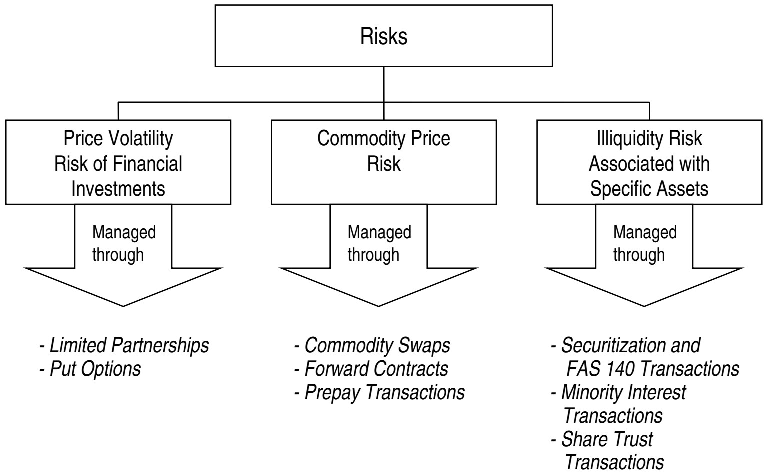


FIGURE 15.3 Risks and Their Management Using Structured Finance

market; commodity price risk can be managed through commodity swaps, forward contracts, and prepay arrangements; and the illiquidity risk associated with specific assets in the firm's portfolio can be managed through securitization and FAS 140 transactions, minority interest transactions, and share trust transactions. This is shown pictorially in Figure 15.3. In what follows, we briefly discuss the underlying economic rationale for each mechanism from a risk-management perspective.

Hedging Price Volatility Risk of Financial Investments

Companies frequently invest in the equity of other companies. Sometimes these are related businesses and sometimes they are not. The objectives of such investments are varied, ranging from passive investments in the hope of a future capital gain to strategic investment with the possibility of acquiring controlling or complete ownership at some future date. For example, Berkshire Hathaway acquired a substantial interest in the insurance company Geico in the 1970s and then watched the value of this investment grow before acquiring the rest of the company in 1995. However, there is a significant risk associated with such investments, and it is the risk that the stock price of the company whose shares the firm in question is holding will fall unexpectedly. Moreover, large holdings are also difficult to sell without moving the price adversely, especially if these are holdings of a firm with thinly traded equity. It may thus be important for the company holding such stock to find mechanisms to hedge its price risk.

One hedge is a *limited partnership*, in which one or more general partners manage the business while "limited" partners contribute capital and share in the profits but take no part in running the business. General partners remain personally liable for partnership debts while limited partners incur no liability with respect to partnership obligations beyond their capital contributions. Death, disability, or withdrawal of a general partner dissolves the partnership unless the partnership agreement specifies otherwise. A company can use a limited partnership as a private equity fund to hedge its security price risk. To see how, suppose company A owns equity in company B and wants to hedge the price risk associated with that equity. It can seek an investment bank's help to set up a partnership in which some third party can become a general partner and the investment bank can become say a limited partner. The general and limited partners can provide capital. Company A can then transfer some of the equity whose price risk it wants to hedge to the partnership, in exchange for which the partnership can provide company A with a promissory note and/or a put option that guarantees that company A can sell back to the partnership its stock in company B for a predetermined fixed price. The partnership can use the capital it has to make other investments. The partnership effectively functions as a private equity fund that uses the capital and other assets it has to sell a put option to company A (in exchange for a portion of company B's stock) that permits company A to hedge its investment in company B.

Derivative securities such as ***put options*** and swaps are also hedging instruments. If the stock in question does not have options trading on it, an investment bank may be able to create a synthetic, nontraded put option whereby a third party may be found to write an option at a negotiated price. Thus, investment banks can play advisory and brokerage roles in helping their clients manage risks by using derivatives, such as options, swaps and futures.

Hedging Commodity Price Risk

Many companies, especially those in the agriculture, food and natural resource industries, are routinely faced with commodity price risks. The prices at which they can purchase their inputs in the future as well as the prices at which they can sell their outputs are uncertain. The risk created by this uncertainty is often hedged, with the help of investment banks. There are various ways in which this is done: commodity swaps, commodity futures and forward contracts, and prepay transactions are three examples.

An alternative to swaps, futures, and forwards is a ***prepay contract***. The basic idea how a prepay contract works is pretty straightforward, although actual contracts can end up looking quite complex. An SPE is first established and the company that wishes to hedge its commodity price risk receives a payment in advance from the SPE in exchange for the company's promise to make future delivery of the commodity to the SPE. This is called a *prepaid forward contract*, which is a form of structured finance. The company selling the commodity in exchange for the advance payment is therefore transferring commodity price risk to the SPE, which in turn passes this risk onto the financiers of the SPE (some of whom may be investment banks). The financiers dissipate their risk through a variety of means, including diversifying across many such investments and hedging using swaps and futures. Investment banks play a key role in this market. They not only help design and set up the structured finance transaction, but may also help establish the SPE. That is, investment banks may provide the funds the SPE uses to make advance payment for the future delivery of the commodity to the SPE. In addition to the prepaid forward contract, the company may also enter into separate derivative contracts with investment banks to hedge the price at which it will acquire the commodity it has promised to deliver to the SPE. Prepaid forward contracts are an important part of structured finance in helping companies manage commodity price risk and create an additional source of financing, i.e., diversify funding sources.

Illiquidity Risk Associated With Specific Assets

Companies are often faced with the problem that many of the assets they own are not very liquid. Hence, if the company needed to change its portfolio mix and divest assets, it would incur substantial costs in converting these illiquid assets into cash. The risks associated with this fall under the heading of illiquidity risk, and this is a risk corporations often wish to hedge, typically with the help of an investment bank. Some of the structured finance transactions that help corporations manage illiquidity risk are described below.

Securitization (also called FAS 140 transactions) are mechanisms by which a corporation can liquefy illiquid assets. Securitization, covered earlier in Chapter 9, is a very popular financing vehicle that deploys a variety of assets like receivables, inventories and the like. The company essentially sells the cash flows of its less liquid assets. Illiquid assets are thus "monetized," and in effect removed from the company's balance sheet. A FAS 140 transaction is designed, with the help of investment banks, to diversify the company's funding sources and "monetize" or "liquefy" assets.

An alternative approach to obtaining financing against an illiquid asset pool is through a *minority interest transaction*. This transaction involves company A setting up a subsidiary in which it has majority interest and which is consolidated with the company for financial reporting purposes. There is also a minority shareholder who has a minority interest in the subsidiary. The minority shareholder is not consolidated with company A for financial reporting purposes. This unconsolidated minority subsidiary generally holds no assets other than a minority interest in the subsidiary owned by company A. The minority shareholder then obtains a mix of debt (typically a loan from some lender) and equity (typically raised from third-party equity investors) financing, with the higher percentage coming from debt. It provides its lenders with a note in exchange for the debt financing. The minority shareholder invests these proceeds from its debt and equity financing in the subsidiary, which then passes the money along to company A. In exchange for this funding, company A transfers to the subsidiary the pool of otherwise illiquid or imperfectly liquid assets it is seeking to "monetize." These could be merchant investments in other natural resource companies and products, receivables, preferred stock of another subsidiary.

Finally, a *share trust transaction* is yet another way investment banks help companies diversify their funding sources and liquefy asset pools. With this arrangement, company A creates a business trust (call it the "issuer") that sells notes and certificates of beneficial interest in the private placement market, typically to institutional investors. The proceeds from these security sales are then transferred to another entity (call it the "holding entity"). The holding entity passes along a portion of the cash to company A to enable it to pay off some debt on its balance sheet or to purchase assets. In exchange for this financing, company A does two things. First, it transfers some assets (the illiquid or imperfectly liquid asset pool it wishes to obtain funding against) to the holding entity. These assets are thus removed from company A's balance sheet; the holding entity and the issuer are not consolidated with company A for financial reporting purposes.

The holding entity takes the cash from the security issuance proceeds that it does not transfer to company A and pools this cash with the assets transferred to it by company A to establish a reserve fund to support payments to the purchasers of the securities issued by the issuer. Moreover, company A establishes a share trust to which it issues preferred stock, issuing enough preferred stock to ensure that it is sufficient to repay the issuer's note when due. Thus, a share trust transaction is another way to use the off-balance sheet financing technique that is an integral part of many structured finance transactions in order to diversify funding sources and securitize illiquid assets.

Conclusion

Investment banks are general purpose financial intermediaries largely in the service of businesses. They perform a wide range of brokerage and asset transformation services, but at base their role is to mobilize financial capital in the service of businesses and other capital users including governments and not-for-profits. Notably, investment banks do *not* provide monetary services as these are the exclusive domain of commercial banks who enjoy the use of governmentally issued liabilities and access to the Federal Reserve discount window.

With the passage of Gramm-Leach-Bliley legislation and the dismantling of Depression- era Glass-Steagall legislation, commercial banks, investment banks, issuance companies have begun a process of consolidation giving rise to spatially and functionally integrated financial intermediaries of enormous size and equally impressive complexity.

Appendix 15.1[9]

IPOs Are Underpriced

The IPO market has generated many research studies, particularly those aimed at explaining why IPOs are consistently underpriced on average, yielding an average return of 16 percent on the first day of issue.[10] Clearly the firms issuing the securities would like to capture some of this gain through higher initial prices. Of course, the danger is that too high a price may discourage investors and the company will not raise the funds needed for investments or restructuring. These two conflicting forces play a role in the IPO models that have been developed.

This addendum will explore three of these models: the information heterogeneity model of why new issues are underpriced, the cascades model explaining how potential investors in an IPO can learn from the purchasing decisions of earlier investors, and the litigation risk/reputation model of underpricing.

9. This material is from "The Value Sphere: The Corporate Executive's Handbook for Creating and Retaining Shareholder Wealth," by John Boquist, Todd Milbourn and Anjan Thakor, VIA Press, 2006.

10. For evidence of the IPO underpricing phenomenon, see Jay Ritter, "Initial Public Offerings," *Contemporary Finance Digest*, Spring 1998, pp. 5–30.

Why New Issues Are Underpriced: Information Heterogeneity

Kevin Rock has proposed an interesting explanation for the underpricing of IPOs.[11] The key to his explanation is that there is a group of investors with limited wealth who are "informed" about the future prospects of the firm raising capital through the IPO, i.e. their information is better than that of other (uninformed) investors, the firm issuing the securities, and its underwriter.

Rock offers two justifications for why informed investors have better information than the issuing firm and its investment banker:

1. All relevant information about the issuing company is disclosed in the prospectus.
2. All the individuals in the market, including competitors and other bankers, collectively know more than the issuer and his investment banker.

Informed investors bid for the issue when the offering price is below the true value, creating the possibility of oversubscription. Likewise, if the true value is below the offering price, the informed investors withdraw and there is the prospect of an excess supply of the new shares. Given the potential for excess demand or excess supply, there is no guarantee that an order for the new shares will be filled. Rather, each investor will be allocated shares in an oversubscribed issue in a random fashion.

With the "all or nothing" participation of the informed investors, an uninformed investor views the probability of receiving an allocation of an overpriced issue as being greater than the probability of receiving an allocation of an underpriced issue. Thus, if issues are correctly priced *on average*, the uninformed investors always buy too much when shares are overpriced and too little when they are underpriced, thereby losing money on average. To compensate for this bias, the issuer must price the shares at a discount to attract the uninformed investors to the offering. And the discount must be just enough to compensate uninformed investors for their informational disadvantage relative to the informed investors. The informed investors still earn more profit than the uninformed. But in effect, the "relative losses" of the uninformed investors are absorbed by the issuing firm, so they at least break even.

Cascades and the IPO Market

Ivo Welch has proposed a cascades explanation for the underpricing of IPOs.[12] His explanation also implies that IPOs will fail or succeed rapidly. Furthermore, Welch demonstrates that it is possible for underpriced offerings to fail and overpriced ones to succeed.

The basic idea is that subscription to an IPO takes place in stages. This means later investors can learn from the earlier investors. Initial sales success implies that the early investors had a favorable view of the new shares, giving impetus to later

11. Kevin Rock, "Why New Issues Are Underpriced," *Journal of Financial Economics*, 15, 1986, pp. 187–212.

12. Ivo Welch, "Sequential Sales, Learning and Cascades," *Journal of Finance*, June 1992, pp. 695–732.

investors to invest. With sufficiently strong initial demand, the information conveyed by the initial subscribers may be so strong that later investors may prefer to disregard their *own* information and subscribe to the issue.[13] This is what Welch calls the "cascade effect," and it can guarantee success if initial investors can be enticed to buy.

By the same token, if initial investors abstain, later investors may stay away as well, even if their own information is favorable. The issue then fails. Thus, there are both negative and positive cascades. Underpricing an IPO is a way to influence early investors to subscribe and increase the probability of success (positive cascade). In the case of the Broadcast.com IPO discussed earlier in this chapter, there was a huge positive cascade.

Litigation Risk, Reputation, and Underpricing

Seha Tinic has suggested that investment bankers underprice IPOs to minimize the probability of being sued by investors if the after-market price declines.[14] Patricia Hughes and Anjan Thakor showed subsequently that litigation risk is neither necessary nor sufficient for IPO underpricing.[15] However, they also showed that there are plausible circumstances in which it is optimal for the investment banker *and* the issuer to underprice because of litigation risk and/or reputational concerns.

The basic idea is that investors can sue the underwriting investment bank and the issuer for misrepresentation if the after-market price falls below the IPO price *and* the post-issue operating financial performance of the firm is below expectations. Such litigation is costly, not only for the obvious legal reasons, but also because it can damage the investment bank's reputation as well as that of the issuer. Underpricing is designed to reduce the probability of an after-market price decline and hence diminish litigation and reputational risks. After all, it is difficult for investors to sue if they make money on a deal. What is interesting is that the extent of these risks—and thus the extent of underpricing—depends on the reputation of the price setter prior to the issue. The better this reputation, the lower these risks. Since investment bankers can be expected to have better reputations than issuers for pricing issues "correctly," an important prediction of the model is that there will be *more* underpricing if the issuer sells stock directly without using an investment banker than if a banker is used. Thus, Jerry probably underpriced less using Alex Butler than he would have if he had tried directly selling to the public on his own.

The IPO as a Branding Event

In the 1990s, IPO underpricing appeared to have been used to gain media publicity for young companies that saw such publicity-generating benefits for their future product or service sales. The Internet IPOs seem to be a good example. Unlike the previous hot IPO sectors like biotech, where start-up companies tried to raise all the

13. Some said that you know the deal is good if Warren Buffett is in.

14. Seha M. Tinic, "Anatomy of Initial Public Offerings of Common Stock," *Journal of Finance* 43, 1988, pp. 789–822.

15. Patricia Hughes and Anjan V. Thakor, "Litigation Risk, Intermediation, and the Underpricing of Initial Public Offerings," *Review of Financial Studies* 5–4, 1992, pp. 709–742.

capital they needed, many Internet companies seem to view the amount of money raised as no more important than being the center of a media buzz.[16]

For example, in the summer of 1998, Broadcast.com told its bankers it would rather stick with an offering price of $18 a share, even though it was apparent it could get a higher price. The *Wall Street Journal*, January 19, 1999, reported,

> We could have had $35 a share," says Chief Executive Todd Wagner. But Broadcast.com 'viewed the IPO as a branding event.' A soaring first-day stock "was a way to launch our name."
>
> It worked for Globe.com as well. Its spectacular opening won a mention in publications as far afield from Wall Street as Sports Illustrated and helped the company attract dozens of advertisers. "We're not one of the random Internet companies any more," exults Mr. Paternot.
>
> Adds the chairman, Mr. Egan: "If I left a few million dollars on the table, so what? The IPO is about getting investment money, but in the case of the Internet, it's also a case of getting a public persona. We wanted to make sure we had a home run, not a double."

Flipping, Spinning, and Recent IPO Developments

IPO underpricing has been viewed by many people as a way for investment banks underwriting IPOs to curry favor with preferred clients or potential clients. It is difficult for retail investors to buy shares in an IPO. Only those customers who have special relationships with either the underwriter or the retail brokerage houses that receive allotments of shares in the IPO are able to purchase the IPO. By underpricing the IPO, the underwriter is able to "reward" these customers who can turn around and sell their shares at handsome profits. The act of buying shares in an IPO and selling them immediately in the after-market (typically on the first day of trading) is called "flipping," whereas the practice of giving IPO shares to favored or potential clients in hopes of winning future business is called "spinning." The SEC has frowned on this practice, particularly if it is seen as a tie-in with sales for future bank services, and has forced banks to allocate some IPO shares to all investors.

Recently, on-line trading has begun to be used as a mechanism to achieve two objectives: (i) permit more retail investors to purchase shares in IPOs, and (ii) reduce IPO underpricing.

For example, Wilt Capital Corp. is a firm that specializes in on-line IPOs. It allocates limited shares of other firms' deals to its own customers on a first-come, first-served basis at the price set by the underwriter. This addresses the goal of broadening retail investors' participation. But it does little to alleviate underpricing.

Hambrecht & Quist LLC goes a step further. It has started a new company that has started selling IPOs over the Internet. The company uses a "Dutch auction" process both to set the offering price and to distribute stock to individual investors.

The plan, called Open IPO, works as follows. Prior to the IPO, potential investors submit bids for the number of shares they would like to buy and at what price.[17] After a few weeks of accepting bids, the IPO offering price is set at the highest price at

16. In fact, many Internet IPOs quickly follow up a successful IPO with another round of financing.

17. They can submit bids as long as they have a brokerage account through W.R. Hambrecht or one of the five small brokerages that have agreed to participate in the process. See Lisa Bransten and Nick Wingfield, "Hambrecht Goes Online for IPOs," *Wall Street Journal*, February 8, 1999.

which all the shares can be sold. Those bidding above the offering price get all of the shares they requested at the offering price. Those bidding at the offering price get a fraction of their orders filled. And those bidding below the offering price get nothing. This is the procedure the major investment banks used in the Google IPO.

The plan has some restrictions. No more than 10 percent of the shares to a single bidder. And Hambrecht has the right to limit the purchases of anyone who wants to buy more than 1 percent of the shares.

Before opening the bidding, Hambrecht sets an expected price range for the stock. This is intended to give investors an idea of what the stock might be worth. Hambrecht has promoted this scheme by ridiculing the current system of IPO distribution as the underwriter telling the lay retail investor, "One for me, none for you."

IPOs, Capital Market Development, and the Average Age of Firms Going Public

IPOs are facilitated by growing investor participation in the public capital market since this provides greater liquidity and potentially higher prices.[18] Given the significantly higher levels of investor participation in recent years, it is not surprising that IPOs have become more popular. Moreover, as a result of this greater popularity, the average age of firms going public has declined from 40 years in 1960 to about 5 years in 2000.[19]

References

Allen, Franklin, The Market for Information and the Origin of Financial Intermediation, *Journal of Financial Intermediation* 1, 1990, 3–30.

Bhattacharya, Sudipto, and Anjan V. Thakor, Contemporary Banking Theory, *Journal of Financial Intermediation* 3, 1993, 2–50.

Boot, Arnoud, Radhakrishnan Gopalan and Anjan Thakor, "The Entrepreneur's Choice Between Private and Public Ownership," *Journal of Finance* 61–62, April 2006, 803–836.

Boquist, John, Todd Milbourn and Anjan V. Thakor, *The Value Sphere: The Corporate Executive's Handbook for Creating and Retaining Shareholder Wealth*, 2006, VIA Press.

Bransten, Lisa, and Nick Wingfield, "Hambrecht Goes Online for IPOs," *Wall Street Journal*, February 8, 1999.

Chemmanur, Thomas, and Kose John, Optimal Incorporation, Structure of Debt Contracts, and Limited Recourse Project Financing, *Journal of Financial Intermediation* 5, 1996, 372–408.

Fink, Jason, Kristin Fink, Gustavo Grullon and James Weston, "IPO Vintage and the Rise of Idiosyncratic Risk," Working Paper, Rice University, 2004.

18. It has been shown that greater investor participation leads to potentially higher stock prices. See Arnoud Boot, Radhakrishnan Gopalan and Anjan Thakor, "The Entrepreneur's Choice Between Private and Public Ownership," *Journal of Finance*, April 2006.

19. See Jason Fink, Kristin Fink, Gustavo Grullon and James Weston, "IPO Vintage and the Rise of Idiosyncratic Risk," Working Paper, Rice University, November 2004.

Garver, Bob, "Equity Party Poopers," *CFO Magazine*, October 2005, 17–19.

Hughes, Patricia, and Anjan V. Thakor, "Litigation Risk, Intermediation, and the Underpricing of Initial Public Offerings," *Review of Financial Studies* 5–4, 1992, 709–742.

Ramakrishnan, Ram, and Anjan V. Thakor, Information Reliability and a Theory of Financial Intermediation, *The Review of Economic Studies*, 1984, 415–432.

Ritter, Jay R., "Initial Public Offerings," *Contemporary Finance Digest*, 5–30, 1998.

Ritter, Jay R., and Ivo Welch, "A Review of IPO Activity, Pricing, and Allocations," *Journal of Finance* 57–4, 2002, 1795–1828.

Rock, Kevin, "Why New Issues Are Underpriced," *Journal of Financial Economics* 15, 1986, 187–212.

Shah, Salman, and Anjan V. Thakor, "Optimal Capital Structure and Project Financing," *The Journal of Economic Theory* 42, 1987, 209–243.

Tinic, Seha M., "Anatomy of Initial Public Offerings of Common Stock," *Journal of Finance* 43, 1998, 789–822.

Thomson Financial, December 31, 2004, Press Release.

United States General Accounting Office, Investment Banks: The Role of Firms and Their Analysts With Enron and Global Crossing, *Report to the Senate Committee on Banking, Housing, and Urban Affairs and the House Committee on Financial Services*, 2003.

Welch, Ivo, "Sequential Sales, Learning and Cascades," *Journal of Finance*, 1992, 695–732.

This page intentionally left blank

PART ♦ IX

The Future

This page intentionally left blank

CHAPTER ♦ 16

The Future

"Rome fell; Babylon fell; Hindhead will have its turn."
George Bernard Shaw

Glossary of Terms

Hedge Funds: Funds that take both long and short positions in a variety of financial instruments to hedge the exposure of the portfolio to unexpected price movements and achieve the highest returns commensurate with the fund's objectives (see Chapter 2).

Universal Bank: A bank that combines commercial banking, investment banking, and possibly insurance and other financial services.

Introduction

Some think that George Bernard Shaw's comment, directed at the British empire, applies to the world's banks.[1] What future does banking have? Banks provide a variety of asset transformation and brokerage services. These services will continue to be demanded and produced, but not necessarily by banks as we know them. To ensure that depositors do not lose faith, governments protect banks in a variety of ways, and therefore monitor and regulate them in order to avoid being exploited.

It does not take a visionary to imagine a world without "traditional banks." But what might the future of banking look like?

1. See *World Banking* in *The Economist*, May 2, 1992.

Our discussion of this issue is organized around three main themes: the likely continuation of the trend toward globalization in banking and its implications, risk management by banks and its implications, and the future of international capital regulation of banks.

Globalization is likely to continue to be an inexorable force since environmental changes are likely to continue to pressure profit margins in traditional banking activities and force banks into new modes of delivering an expanded array of financial services. Part of the new opportunities will be provided by universal banking and part by financial innovations and new products, but part will come from the increasing presence of U.S., European, and Japanese banks in emerging markets, i.e. through greater globalization.

Universal banking, financial innovation and grater forays into emerging markets will all create additional risks for banks, which will compel regulators to rethink the design of banks and the scope of the public safety net.

As part of this rethinking, it is likely that the Basel capital regulations will also evolve. There is much to expect in the future.

Future Opportunities for Banks: Expanded Role for Relationship Banking and the Implications for Universal Banking, Financial Innovation, and Globalization

As banks' profit margins are squeezed by competition, they will have to keep searching for new opportunities. An attractive source of such opportunities is likely to be deeper relationships with existing and new customers.[2] As we saw in Chapters 5–7, relationship banking can be profitable for banks. The question is: How can banks deepen their relationships and increase the value added for their customers by these relationships?

One way to increase the value added in relationship banking is to engage in financial innovations, creating new commercial and investment banking products that help banks' customers to diversify funding sources, improve liquidity and lower the costs of various sources of capital. Off-balance sheet financing (Chapter 8) and securitization (Chapter 9) are good examples of the kinds of innovations that have helped banks' customers achieve these objectives and generated profits for banks.

In the future, the innovation possibilities for banks will be even greater because of the dismantling of the Glass-Steagall Act, which now provides banks with the opportunity to combine traditional insurance and commercial and investment banking products to create new, unconventional products. Thus, universal banking may well open the door to the introduction of hitherto-unfamiliar financial products.

Besides innovation and universal banking, globalization will provide another avenue for banks to deepen relationship banking. Increasingly the customers of banks are seeking growth opportunities in a variety of global markets. As these customers become more global, the value that global banks can provide for these global customers also becomes potentially greater. It is because of this that in 2004

2. Boot and Thakor (1997) show that as interbank competition increases, a potentially successful approach for banks is to increase the value added in relationship banking.

Citigroup became the largest underwriter of corporate debt, not just in the United States, but around the world.[3]

In Europe, Morgan Stanley was the largest underwriter of equity and equity-related offerings in 2004. Bank of America and Deutsche Bank provided the cash management, investment banking, credit and capital-raising needs of their multinational customers around the globe. Some banks rely on partnerships with nonaffiliated local banks to provide global financial services. The advantages of providing a comprehensive range of financial services comes from cross-sectional and intertemporal information reusability (see Chapters 2 and 3).

While being global almost never guarantees exclusive access to a customer's global business—there are always specialized local banks that can be formidable competitors for some types of services—a CFO survey found that 43 percent of their respondents with overseas banking needs stated that global banks could meet all of their needs.[4] An area in which global banks appear to have a particular advantage is cross-border mergers and acquisitions. The advantage comes from the fact that the bank can make a deal for a client in one country and distribute or sell that deal into another market, thereby helping the client diversify its funding sources. Table 16.1 provides information about the major global banks in 2005.

Emerging markets, while a potential fountain of growth opportunity for global banks, still present a big challenge. Part of the reason is Byzantine banking regulations in many emerging markets that can impede global banks.

In other instances, there are outright restrictions on the types of services that can be offered by banks or restrictions on foreign ownership of local banks. India and China, the two biggest emerging markets, have proved particularly challenging, and no bank has made major headway yet in either market. Indian banking regulations effectively require multinational companies to use local banks for some services. The Chinese government limits the degree to which foreign banks can participate in local currency operations. However, it is likely that things will change in the future.

TABLE 16.1 Major Global Banks in 2005

Citigroup	Greatest global reach of any bank; $1.5 trillion in assets; strongest emerging market presence; offices in over 100 countries; major player in commercial and investment banking.
Hong Kong & Shanghai Bank (HSBC)	$1.5 trillion in assets; offices in almost 80 countries; strong presence in Asia; strong globally in commercial banking, but not top 10 in global underwriting.
JP Morgan Chase	$1.2 trillion in assets; does business in 50 countries; fifth-largest underwriter.
Deutsche Bank	$1.2 trillion in assets; operations in 74 countries; established presence in U.S. by acquiring Bankers Trust in 1999.
Bank of America	$1.2 trillion in assets; offices in 31 countries.

Source: Myers (2005a).

3. See Myers (2005a). He quotes Brown-Forman CEO Phoebe Wood, "Global banks are powerhouses in raising capital."

4. See Myers (2005a).

Risk Management by Banks

One of the areas in which things have perhaps changed the most is risk management. As banks offer an increasing array of services, they are also exposed to an increasing variety of risks. These risks arise because banks are increasingly relying on derivatives and various types of off-balance sheet transactions.

Banks are managing these risks through swaps, derivatives and other contingent claims. They also package their subordinate debt into various forms of structured finance securities (recall Chapter 15) and sell these to investors, with hedge funds often buying the riskiest securities. This is referred to as a "risk dispersal," which is a process by which banking risk is dispersed throughout the financial market. A survey by Greenwich Associates found that hedge funds controlled as much as 30 percent of the trading volume in high-yield bonds and 26 percent of leveraged loans in the U.S. In the structured finance market, hedge funds play a critical role in financing the least liquid, highest-yielding subordinate tranches of transactions.

All of this raises questions about whether the capital markets are well equipped to process and absorb these risks. The collapse of Long Term Capital Management (LTCM) has heightened these concerns; see Chapter 13. The Counterparty Risk Management Policy Group (CRMPG) II noted that improvements had been made in capital market risk processing, but warned that operational risks resulting from the rapid rise in the use of credit derivatives and other financial innovations since LTCM's decline could pose future threats. In June 2005, the Bank for International Settlements (BIS) cautioned that structured finance, especially structured credit derivatives, has grown so complex that market participants are compelled to rely heavily on ratings to comprehend the risks involved. Yet, the partitioning of overall risks into a variety of pieces skews rating methodologies and raises questions about the reliability of ratings. This raises the specter of systematic risk and has resulted in a variety of reports on the subject. This information is summarized in Table 16.2.

The Basel Initiative and Future Capital Accords

In earlier chapters we discussed the Basel II Capital Accord. This capital regulation is likely to affect the future evolution of banks as the three pillars of Basel II will bring a combination of forces to bear on banks: regulatory capital requirements, regulatory monitoring and market discipline. However, unlike the other nine countries participating in Basel II, the United States is not requiring all of its banks to comply with it, although another 18 or so banks are expected to do so voluntarily. See Table 16.3.

In addition, unlike other countries, U.S. regulators are requiring U.S. banks to adopt only the "advanced" Basel II model, which is the most sophisticated and complex version of the accord. Banks in other countries can choose between the advanced model, the "foundation" model, which is more prescriptive in calculating risks, and the "standardized" model, which affords the least flexibility in calculating minimum capital requirements (see Chapter 12 for a discussion of Basel II).

One concern among smaller U.S. banks is that they will be placed at a disadvantage because Basel II permits larger banks to keep lower capital. Consequently, U.S. regulators are considering a refinement of Basel I for smaller banks that is less complex than Basel II, but still generates more risk-sensitive capital requirements than Basel I.

TABLE 16.2 Reports About Systematic Risk

Date	Document	Issued by	Comments
May 14, 2004	"Interagency Statement on Sound Practices Concerning Complex Structured Finance Activities"	Office of the Comptroller of the Currency, Office of Thrift Supervision, the Federal Reserve, the FDIC, and the SEC	Proposes ways for banks to monitor structured-finance transactions set up for corporate clients to avoid legal and reputational risks (to the banks). The banking industry responds with strong objections. A revision is still expected.
May 5, 2005	"Risk Transfer and Financial Stability"	Federal Reserve Chairman Alan Greenspan	Seeming to hedge, though not reverse, his previous praise for derivatives, Greenspan says that regulators can't track credit risk transferred outside the banking system (typically to hedge funds), that risk-mitigation benefits of derivatives could be undermined by contractual failures, and that some market participants are not exercising sufficient care.
May 12, 2005	"Remarks Before the Foundation Financial Officers Group"	SEC Chairman William Donaldson	In remarks on "staving off future crises," Donaldson defends requiring hedge-fund registration, citing the size of the market and the difficulty of detecting fraud.
June 1, 2005	"Operations Benchmarking and FpML Use Survey"	International Swaps and Derivatives Association	Citing "advances," the ISDA reports the average backlog for credit-derivative confirmations is 11.6 days. Only 40 percent of such contracts generate automated confirmations.
June 10, 2005	Quarterly Review: "Structured Finance: Complexity, Risk, and the Use of Ratings"	Bank for International Settlements	Notes that the complexity of structured-finance products makes credit ratings less reliable, "even as their complexity creates incentives to rely more heavily on ratings."
June 15, 2005	"Report and Recommendations Pursuant to Section 401(c) of the Sarbanes-Oxley Act of 2002 on Arrangements With Off-Balance Sheet Implications, SPEs, and Transparency of Filings by Issuers"	SEC	Although inspired by Enron-style complex structured transactions, the SEC's off-balance sheet report also focuses on more mundane off-balance sheet issues, such as leasing and pension-fund accounting. Its discussion of leasing, however, takes a swipe at the market's response to Fin 46(R), warning that the use of securitizations and derivatives in accounting-motivated transactions may already be undermining FASB's effort.

Date	Document	Issued by	Comments
June 27, 2005	Annual report	Bank for International Settlements	Warns that risks involved in structured-credit products may not be fully understood by all market participants. Reiterates that credit ratings may be misleading, adding that standard portfolio risk models may also be inadequate. Also warns that credit conditions fostering growth of CDS and CDO markets may not continue, and that it remains to be seen how these "markets would handle a string of credit blow-ups or a sharp turn in the credit cycle."
July 18,2005	"Hedge Funds: An Emerging Force in the Global Credit Markets"	Fitch Ratings	Warns that the impact of hedge funds on credit markets is poorly understood and that a forced deleveraging could be felt across multiple credit-market segments, reducing access to high-yield debt and structured-finance securities.
July 27, 2005	"Toward Greater Financial Stability: A Private Sector Perspective"	Counterparty Risk Management Policy Group II	Issued by the same group that autopsied Long-Term Capital Management in 1999, this report suggests that operational risks pose the greatest current threat to the financial system. It also warns banks and others to make sure investors—including corporations—understand the derivatives they buy.
August 2005	"Don't Bank on Strong Governance: Observations on Corporate Governance in U.S. Banks"	Moody's	Notes that corporate governance is an increasingly important consideration in bank ratings. Among other conclusions, notes that Basel II's requirement to identify and mitigate risk—including operational risk—is "becoming harder as the pace of innovation in financial instruments quickens."
August 2005	"Systemic Risk and Hedge Funds"	Nicholas Chan et al., MIT	Concludes that "systemic risk is currently on the rise" for hedge funds and that their "symbiotic relationship with the banking sector" means that hedge-fund risk exposures "may have a material impact on the banking sector."

Source: Reason (2005).

TABLE 16.3 The Basel Banks?

Banking regulators haven't specified which U.S. banks and thrifts, or even how many, must comply with Basel II. Their guidelines suggest that roughly the first nine institutions on this list must comply, while the rest are likely to adopt it voluntarily.

Bank	**Assets (in thousands)**
1 Citigroup	$1,547,789,000
2 Bank of America	1,251,037,147
3 JPMorgan Chase	1,171,283,000
4 Wachovia	511,840,000
5 Wells Fargo	434,981,000
6 HSBC North America Holdings	372,555,243
7 Taunus (Deutsche Bank)	366,293,000
8 Washington Mutual	333,742,732
9 U.S. Bancorp	203,981,000
10 Suntrust Banks	168,952,575
11 Countrywide Financial	158,617,821
12 Citizens Financial Group	148,491,012
13 ABN AMRO North America Holding	145,024,570
14 National City	143,975,359
15 Golden West Financial	112,587,849
16 BB&T	105,835,324
17 State Street	104,275,118
18 Fifth Third Bancorp	103,159,676
19 Bank of New York	103,110,000
20 Keycorp	91,010,081

Source: Federal Reserve, SNL Financial. Thrift data as of March; bank holding company data as of June 30 and Myers (2005b).

However, it is an open question whether the concerns of smaller U.S. banks are well founded. Most developed countries require banks to hold 6 percent or more of Tier-I capital in order to qualify as "well capitalized." As a consequence, most banks are likely to stay above this level, with the average ratio likely to be around 8 percent.

How will Basel II affect corporate lending banks? Investment-grade corporate borrowers are likely to find commercial loans easier to obtain and less expensive as well. Basel I did not distinguish among corporate loans of different risks, but Basel II does. This may make lower-risk corporate borrowers more attractive to banks. Moreover, secured bank loans—on which banks have the flexibility to keep lower capital—may replace commercial paper for some borrowers. Basel II may also encourage banks to make greater use of risk-mitigation tools like guarantees, portfolio insurance, and credit derivatives. 364-day loan commitments may decline or become more expensive for borrowers since these required no capital under Basel I but do require capital now.

An interesting question is whether one would have a need for such elaborate capital requirements if the deposit insurance safety net was scaled back or eliminated. Our view, expressed in earlier chapters, is that the answer is no, but this is a hard question to answer definitively. It seems reasonable to postulate, however, that as long as regulators remain concerned about the confidence of financial market participants in the soundness of the banking system, capital regulation is unlikely to go away altogether. But somewhat paradoxically, as risk management becomes more complex, regulators may find it more difficult to implement risk-based capital requirements due to limited ability of regulators to fully comprehend the various risks banks will be exposed to. It is likely that relying more on the discretion of banks to adjust their capital levels to the risks they take on, as well as the discipline of the capital market, may become an increasingly significant imperative for regulators.

Conclusion

The main message of the chapter, and an important message of this book, has been that banks are at their best when they use liquid deposits to finance illiquid assets that carry default risk and require expert monitoring. Performing this asset transformation carries with it credit, liquidity, interest rate, operational, reputational and other risks. Advances in information processing technology have helped banks to manage these risks better. Credit risk is managed with expert systems and other credit scoring models that supplement human judgment. Liquidity risk is managed by liquefying asset portfolios through securitization. Interest-rate risk is managed with an ever-expanding arsenal of derivatives including swaps, futures, and options. And the assortment of other risks are managed through an *enterprise risk management (ERM)* approach, with sufficient capital to absorb the risks.

It is possible that numerous hard-to-predict changes will occur in this century. As we look ahead, we can expect the boundaries between financial institutions and markets to disappear, as virtually all assets will become liquefied. Financial transactions will involve global trading of claims against all kinds of assets, with information about these assets globally available electronically at the click of a mouse. Regulatory restrictions on banks will becoming increasingly untenable in this age of frictionless global trading. Banks, as we know them, will cease to exist. Financial intermediaries, however, will play an even more important role in liquefying assets, mobilizing capital, and pricing and managing risk.

Review Questions

1. How do you believe banks will evolve in the future? Will they become more universal or more narrow in scope?
2. How do you believe risk management in banks will evolve?

References

Boot, Arnoud W.A., and Thakor, Anjan V., "Banking Scope and Financial Innovation," *Review of Financial Studies*, 10–4, 1997, 1099–1131.

Myers, Randy, "How Global Is Your Bank?" *CFO, Banking & Finance Special Issue*, October 2005a, 24–30.

——, "Regulation: Basel Faulty?" *CFO, Banking & Finance Special Issue*, October 2005b, 51–53.

Reason, Tim, "Who's Holding the Bag?" *CFO, Banking & Finance Special Issue*, October 2005, 40–46.

Index